1st	Edition	Published	.. ..	1921
2nd	,,	,,	.. ..	1923
3rd	,,	,,	.. ..	1925
4th	,,	,,	.. ..	1927
5th	,,	,,	.. ..	1929
6th	,,	,,	.. ..	1934
7th	,,	,,	.. ..	1936
8th	,,	,,	.. ..	1939
9th	,,	,,	.. ..	1949

STATISTICS AND THEIR APPLICATION TO COMMERCE

NINTH EDITION

BY

A. LESTER BODDINGTON

WITH CHAPTERS ON INDEX NUMBERS

BY

A J. H MORRELL, M.A

H.F.L. (PUBLISHERS) LTD.
29, MINCING LANE
LONDON, E.C.3.
[1949.]

PUBLISHER'S NOTE

THE continuous demand for "Statistics and their Application to Commerce" has prompted the publishers to issue a new edition with more detailed information on Index Numbers.

The Chapters on Index Numbers have been prepared by Mr. A. J. H. Morrell, M.A. Chapter IX deals with the general theory of index numbers and supposes a moderate acquaintance with algebra, such as most students of economics, statistics or accountancy should possess. No apology is made for this, but readers who lack the mathematical equipment to understand the whole of this part will, it is hoped, follow Chapter X dealing with index numbers in common use. This does not attempt to be exhaustive, but describes some of the more important official and other index numbers in current use.

Acknowledgments are due to the Controller of H.M. Stationery Office, the Editor of the *Economist*, the Royal Statistical Society and the London and Cambridge Economic Service, for their kind permission to reproduce copyright matter, and to various people who have made helpful suggestions, and also to Messrs. Powers-Samas Accounting Machines, Ltd., for providing the blocks for the illustrations appearing in the Chapter on Mechanical Aids.

H.F.L. (PUBLISHERS) LTD.

January, 1949.

TABLE OF CONTENTS

	Page
PUBLISHER'S NOTE	v
LIST OF TABLES	xii
LIST OF ILLUSTRATIONS	xiii
INTRODUCTION	xiv

PART I

COMPILATION OF STATISTICAL DATA

CHAPTER I

BUSINESS STATISTICS AND THEIR IMPORTANCE

§ 1.	Estimates and Their Uses	3
2.	The Budget and Its Compilation	5
3.	Object of Statistics	6
4.	Statistics in Commerce	7

CHAPTER II

COMPILATION AND USES OF STATISTICS

§ 1.	The Unit of Data to be Collected	11
2.	Homogeneity of the Data	12
3.	Completeness of the Data	16
4.	Stability of the Data	18
5.	Representative Data	19
6.	Methods of Sampling	21
7.	Methods of Collecting Data	22
	(a) Direct Method	23
	(b) The Questionnaire	24
	(c) Published Data	26
8.	Fallacious Comparisons	28
9.	Units for Comparison	31
10	Commercial Uses of Statistics	40
11.	Divisions of Statistical Data	42
12.	Summary	43

CHAPTER III

ACCURACY AND APPROXIMATION

		Page
§ 1.	Accuracy	46
	(a) Standard of Accuracy	47
	(b) How Degrees of Accuracy may be Shown	48
2.	Approximation	48
	(a) Methods of Approximation	48
	i. Approximation to nearest number or specified decimal place	48
	ii. Approximation by discarding figures	51
	iii. Approximation by adding figures	51
	(b) Results of Approximation upon Percentages	52
3.	Errors	53
	(a) Measurement of Errors	53
	i. Absolutely	53
	ii. Relatively	54
	(b) Classes of Error	54
	i. Unbiassed	54
	ii. Biassed	55
	(c) Estimating the Error	56
	i. When the Error is Unbiassed	56
	ii. When the Error is Biassed	58
	iii In an Estimated Average	59
4.	Fallacies in the Use of Percentages	60
5.	Ratios and Their Use	64
6.	Effect of Errors on Ratios	66
7.	Decimals and Their Advantages	67

CHAPTER IV

TYPES AND AVERAGES

§ 1.	Use of Averages	69
2.	The Arithmetic Average	70
	(a) The Simple Arithmetic Average	70
	i. Short Method of Computing	70
	ii. Advantages of the Simple Arithmetic Average	71
	iii. Disadvantages of the Simple Arithmetic Average	72
	iv. Commercial Uses of the Simple Arithmetic Average	75
	v. Moving Average Standard	77
	(1) Uses of Varying Average Standards	77
	(2) Short Term Fluctuations	80
	(3) Uses of the Trend	80
	vi. Progressive Average	81
	(b) The Weighted Arithmetic Average	83
	i. Definition of the Weighted Arithmetic Average	84
	ii. The Weights to be Used	85
	iii. When the Weighted Average must be Used	87
	(1) When Classes of the same Group contain varying number of items	87
	(2) When proportions change	89
	iv. Commercial Uses of the Weighted Arithmetic Average	90
	(c) The Descriptive Average	92
3.	The Geometric Average	94

CHAPTER IV (*continued*)

		Page
§ 4.	The Median	96
	(*a*) Arraying the Data	96
	(*b*) Frequency Tables	96
	(*c*) Rule for Calculating the magnitude of the Median	98
	(*d*) Quartiles	98
	(*e*) Deciles and Percentiles	99
	(*f*) Advantages of the Median	99
	(*g*) Disadvantages of the Median	100
	(*h*) Uses of the Median	100
5.	The Mode	101
	(*a*) Method of locating the Mode	102
	(*b*) Advantages of the Mode	107
	(*c*) Disadvantages of the Mode	107
	(*d*) Commercial Uses of the Mode	108

PART II

PRESENTATION AND USES OF STATISTICAL DATA

CHAPTER V

TABULATION

§ 1.	Importance of Proper Tabulation	114
2.	Kinds of Tabulation	116
	(*a*) Simple	116
	(*b*) Complex	116
3.	Rules for Tabulation	118
4.	Data to be Tabulated	121
5.	Title of the Table	123
6.	The Uses of Percentages and Averages in Tabular Work	128
7.	Tabulation of Commercial Data	136

CHAPTER VI

THE GRAPHIC METHOD

§ 1.	Application of the Graphic Method	145
2.	The Frequency Polygon	146
3.	The Normal Frequency Curve	149
	(*a*) Skewness	150
	(*b*) Smoothing the Curve	153
	(*c*) When Data consists of items of known magnitude	154
4.	The Ogive	157
5.	Kinds of Diagrams	159
	(*a*) The "Block" Diagram	161
	(*b*) The "Bar" Diagram	162
	(*c*) The Graph	162
	i. The Natural Scale Graph	166
	ii. The Logarithmic Scale Graph	167
	(*d*) The "Bar" Diagram to show Components	171
	(*e*) Comparisons of Areas and Volumes	174
	(*f*) Diagrams to show Three Factors	177
	(*g*) Diagram showing Two Scales	180
6.	Advantages and Disadvantages of Graphs	181
7.	Rules for Constructing Diagrams	182

CHAPTER VII

THE GRAPHIC METHOD APPLIED TO COMMERCE

		Page
§ 1.	Records of Sales	187
2.	Utility of the Moving Average Line	190
3	Comparisons of Quantities and Values	194
4.	Purchases and Sales	196
5.	Price Fluctuations	199
6.	Gross Profit and Expenses	203
7.	Turnover and Net Profit	208
8.	Departmental Records	210
9.	Financial and Control Records	211
10.	Factory Records	213
11.	Cost Records	215
12.	Transport Records	217

CHAPTER VIII

METHODS OF COMPARISON

§ 1.	Dispersion	220
	(a) Measuring Dispersion	221
	i. The Range	221
	ii. Absolute and Relative Dispersion	223
	iii. The Moments of Dispersion	223
	(1) The First Moment	223
	(2) The Second Moment	227
	(3) The Third Moment	230
	(4) The Modulus	231
	(5) The Quartile Measure and Co-efficient of Dispersion	231
	(6) The Lorenz Curve	233
2.	Skewness	236
	(a) The First Measure and Co-efficient	236
	(b) The Second Measure and Co-efficient	237
	(c) The Third Measure and Co-efficient	239

CHAPTER IX

INDEX NUMBERS IN GENERAL

§ 1.	The Nature and Purpose of Index Numbers	241
2.	The Limitations of Index Numbers	242
3.	The Relative	243
4.	The Arithmetic Mean (A.M.)	245
5.	The Geometric Mean (G.M.)	246
6.	The A.M. and G.M. Compared	250
7.	The Chain Base Method	252
8.	The Weighting System	254
9.	The Aggregative Method	255
10.	Changing the Base Year	258
11.	Index Numbers of Quantity	260

CHAPTER X

INDEX NUMBERS IN COMMON USE

		Page
§ 1	Board of Trade Wholesale Price Index	265
2.	"The Statist" Wholesale Price Index Numbers	267
3.	"The Economist" Index of Wholesale Prices	268
4	"The Economist" Sensitive Price Index	270
5.	Import and Export Price Indices .	271
6.	Volume and Value of Imports and Exports .	274
7.	Cost of Living Index Number	275
8.	Interim Index of Retail Prices	277
9.	Interim Index of Industrial Production ..	278
10	Index of Industrial Production (London and Cambridge Economic Service) . .	280
11.	Rates of Wages Index Number .	282
12.	Prof. Bowley's Wage-Rates Index Number	283
13.	Various Indices of Sales and Stocks .	285
	(a) Index of Retail Sales	285
	(b) Index of Sales of Independent Retailers .	285
	(c) Trade of Wholesale Textile Houses . ..	286
	(d) Wholesale Stocks and Sales of Clothing	286

CHAPTER XI

CORRELATION

§ 1.	Definition of Correlation	288
2.	Correlation of two Factors	288
3.	Co-efficient of Correlation	292
	(a) For Long Term Fluctuations	292
	(b) For Short Term Fluctuations	294
	(c) The Probability of Error	294
4.	Ratio of Variation	295

CHAPTER XII

THE CENSUS AS AN AID TO SCIENTIFIC BUSINESS

§ 1.	Information obtainable from the Census	301
2.	Value of Census Statistics to Business Men	303
3.	Density of Population and Market Possibilities	304
4.	*Per Capita* Consumption of Commodities	305
5.	Estimating the Quality of Markets	306
6.	Occupational Lists and Labour Supply	307
7.	Localisation of Industries	309
8.	Value of Census Reports to Producers and Manufacturers ..	311

CHAPTER XIII

BUSINESS RESEARCH.

§ 1.	Profit Earning Possibilities of Research Work	315
2.	Co-ordination of Producing and Selling Organisation	316
3.	General Functions and Objects of Business Research	319
4.	Qualifications of a Business Research Worker	320
5.	Important Statistical Comparisons	321
6.	The Organisation of Production	322
7.	The Problems of Selling	324
8.	Studying Competitors	327
9.	Other Factors to be studied	327

CHAPTER XIV

MECHANICAL AIDS TO STATISTICAL WORK

§ 1.	Calculating Devices	331
2.	The Powers Sorting and Tabulating Machinery	332

LIST OF TABLES

		Page
A.	Total Gross Receipts and Total Working Expenditure of Railways in the United Kingdom (Actual Figures)	50
B.	Total Gross Receipts and Total Working Expenditure of Railways in the United Kingdom (Approximated to Nearest £1,000)..	51
C.	Total Gross Receipts and Total Working Expenditure of Railways in the United Kingdom (Approximated to Nearest £1,000,000)	53
D.	Table showing the Results of Biassed and Unbiassed Errors ..	55
E.	Table showing a Short Method of computing the Arithmetic Average	71
F.	Ratio Turnover bears to Standards calculated on Varying Bases..	79
G.	Table illustrating the Application of the Progressive Average ..	90
H.	Number and Membership of Co-operative Societies marketing Member's produce for Eleven Years .	116
J.	Failures recorded in United Kingdom in Nine Years	117
K.	Engine Hours "In Traffic" of the Engines of Railway Companies during a Month	119
L	Freight Train Traffic on British Railways for a Month . ..	122
M.	Analysis of the Bank of England Return for the Fourteen Weeks ended January 26, 1927	124
N.	Analysis of the Bank of England Return showing Improvements in the Tabulation	125
O.	Tonnage of Freight Traffic on the Railways of England and Wales for a Month	127
P.	Table showing the Frequency with which certain Data occurs ..	146
Q.	Table showing the Frequency with which certain Data occurs when Skewness is present	150
R.	Table showing the Frequency and Cumulative Frequency of Data	155
S.	Table showing the Monthly and Accumulated Sales for Each of Two Years	187
T.	Table showing the Turnover of a Business in Quantity and Value..	195
U.	Table showing the Monthly Purchases and Sales for One Year..	197
V.	Average Price of British Wheat in Each Month for Ten Years ..	201
W.	Table showing the Expenses, Gross and Net Profit of a Business for Ten Years	205
X.	Table showing the Number of Articles sold, Gross Profit, Standing Charges and Net Profit	208
Y.	Table showing the Deviations of the Average Indices from the Index Numbers for each of Thirty Years	290

LIST OF ILLUSTRATIONS

	Page
Specimen Questionnaire	25
Diagram showing a Series of Eighty Items arrayed in order of Magnitude	142
Curve showing an Array of Eighty Items	144
Diagram showing Frequency when Items are Grouped	147
Diagram showing Frequency with which Items occur when Skewness is present	151
Diagram showing Frequency with which Items occur	156
Ogive to show Cumulative Frequency	158
Block Diagram showing Annual Turnover of a Business for Twenty Years	160
Bar Diagram to show the Same Data as in Diagram No. 6	163
Curve on Natural Scale showing the Same Data as in Diagram No. 6 ..	165
Curve on the Logarithmic Scale to show the Same Data as in Diagram No. 6	169
Bar Diagrams showing the Banker's Clearing House Returns for Two Successive Years	173
Comparisons of Squares on the Basis of Area and of Length of Base ..	175
Comparison of Circles on the Basis of Area and of Radii	176
Circular Diagram showing the composition of the Banker's Clearing House Returns for One Year	178
Rectangles used to show Three Factors	179
Diagram showing Monthly and Accumulated Sales for Each of Two Years	188
Diagram showing Comparison of Quantities and Values *facing*	196
Diagram showing Purchases and Sales for Each of Twelve Months ..	198
Diagram showing the Ten-Year Monthly Average Price of English Wheat	202
Diagram showing Gross Profit, Expenses and Net Profit for Ten Years	206
Diagram showing Monthly Gross and Net Profit and standing Charges..	209
Lorenz Curve showing distribution of Profits in two groups	234
Diagram showing the Short Term Fluctuations in Supply and Price *facing*	288
Galton Graph	298
Specimen of Card used in connection with the Powers Machinery ..	333
Powers Automatic Key Punch	334
Powers Sorting Machine	335
Powers Tabulating Machine	336
Specimen of Tabulation on Powers Machine	338

INTRODUCTION

THE business man of to-day is met by many problems which did not confront his predecessors, and chief among them is the tendency to limit by artificial means the rate of profit which he may earn. In order, therefore, to show increasing profits from his business he is compelled either to (*a*) increase his turnover, or (*b*) reduce his expenses and eliminate waste of every kind. The former is, of course, possible, but the ever increasing competition often means that additional turnover is only obtained by increasing the expenditure to an extent which makes the return on such increased turnover much lower than that normally obtained. The latter is, therefore, of great importance, and applies even in those cases where turnover is increasing—in short, Scientific Management is to-day the only hope of showing better profits. Waste, whether of material and/or labour must be eliminated, preventable losses must not take place, the law of increasing returns applied to its fullest extent, and efficient machinery must be introduced to save both time and labour wherever possible. Operating costs, distributing costs, and general expenses must all be watched most closely, since an increase in any of these components may result in the profits being swept away, and a good year's business show a loss.

Up-to-date machinery and time-saving methods are being gradually introduced throughout our offices, but many heads of business houses are apt to regard the purchase price of such machinery and appliances

as an expenditure, whereas it is in reality a first-class investment. This point of view is probably the cause of our business men being regarded as conservative in their methods.

Business is a science, and the office may therefore be regarded as the laboratory, and as such needs to be efficiently fitted to carry out the work for which it is intended. In order to succeed in any business to-day, the business man must study all the factors which enter into the production, buying and selling, exporting and importing of the goods in which he deals. He must know the right time to buy, the right price to pay, the quantity to buy, and the correct place to buy. He must know when to place the article upon the market, the price he can obtain, and the people to whom it is to be sold. He must study transport facilities, and investigate and compare the costs of the various methods of transport which may be adopted, not overlooking the importance and convenience of speedy and certain deliveries. He must know how to create a market for a new product, or how to stimulate the demand for his goods. Financial conditions need close and careful study in order that he may know when it is advisable to expand and when to contract credit, or when it may profit him to purchase his goods and pay for them by raising a loan. If he be an importer or exporter, he must know the influences which affect foreign exchanges so that he may pay or collect his debts at the most propitious moment. He must know when he may speculate with a reasonable expectation of profit, and must be able to come to an immediate decision with regard to any proposition which may be placed before him. This calls for efficiency of a very high order.

The time has come when old-fashioned methods must be scrapped, old ideas swept away, and new and improved methods used in order that business may expand, profits increase, and development be rendered possible.

During the past few years we have seen great advances made in business organisation and methods. Where formerly the personal influence of the proprietor attracted business, the present tendency is of an impersonal nature. Buyers purchase their commodities in the best market, irrespective of the person with whom they are dealing. Small businesses, where the proprietors knew their regular customers, are passing away. Multiple shops and Departmental Stores are springing up. Amalgamations are rife and large businesses are increasing in number. As a consequence of this tendency the principal must rely upon others for the specialised and detailed work necessary, and can then devote his energies to directing the general policy of the business. He can be greatly aided by being supplied with statistical records, which will enable him to analyse the situation and obtain a panoramic view of the business as a whole. Thus he will supplement his own abilities by making use of scientific methods, will obtáin the true perspective of the activities of his subordinates, and gauge the possibilities of profit-making opportunities, and so be materially assisted in building a successful and profitable business.

STATISTICS

AND

THEIR APPLICATION

TO COMMERCE.

PART I.

THE COMPILATION

OF

STATISTICAL DATA.

SYNOPSIS TO CHAPTER I.

BUSINESS STATISTICS AND THEIR IMPORTANCE.

§ 1.—Estimates and Their Uses.

2.—The Budget and Its Compilation.

3.—Object of Statistics.

4 —Statistics in Commerce.

CHAPTER I.

BUSINESS STATISTICS AND THEIR IMPORTANCE.

Every business man finds it necessary to keep a record of his transactions, but frequently these records consist only of those financial aspects of his activities which enable him to prepare proper accounts periodically, consequently the advantages which accrue to those who use information as to quantities, qualities, etc., accumulated from time to time, for placing their business affairs upon a scientific basis, are not obtained. It is only when we compare items one with another that we use statistical methods.

§ 1.—Estimates and Their Uses.

Financial Institutions such as Banks and Insurance Companies can only carry on their businesses by the intelligent utilisation of past experience. As a result of studying previous records a Banker knows what cash reserve it is necessary to maintain in order to meet normal calls made upon him, and moreover can vary his Reserves as the need of the occasion demands. In an agricultural district, for example, it may be necessary to increase the amount of cash in hand at harvest time, in order to meet the demand of farmers

for money wherewith to pay the wages of the additional men required at this time. Insurance Companies base their premiums upon the results of the risks covered over a large number of years, or from other data at their disposal. Railway Companies, in order to provide for the great crowds which habitually travel at holiday times, must estimate the number of persons they may be called upon to carry, and the distance and direction they will travel. These estimates are naturally based upon what has happened in the past, and thus we find additional trains provided on some services, and a curtailed service on other sections of the line. While it is true that the estimates may not be, and usually are not, correct, the system of utilising past experience has much to recommend it, since the error will be in large measure due to factors outside the control of the company, such as bad weather or a change in popular taste. In the United States of America there are many operators in "futures." They base their speculative operations on the possible crops of any particular commodity in which they are interested, and their estimates of the quantity which will be available are obtained by taking the average yield per acre over a large number of years, and multiplying by the number of acres which it is estimated are under cultivation. The Government publishes much information as to yield, while the acreage is estimated by the commercial bureaux and newspapers from the information collected by their correspondents. Such estimates, like those of the Railways, may be incorrect owing to the fact that natural forces come into play, but the estimates of the yield will be found to be approximately correct in a year which is an "average" one, that is, one in which no abnormal influences affect the crops. If

natural factors could be eliminated errors would almost disappear. To take an example, the world's output of gold can be estimated almost exactly, since the only factor in normal circumstances which will affect the output to any extent is the discovery of a new gold-field. In recent times, however, the increase in the price of gold has caused hoarded gold to be thrown on the markets, but such a position is clearly abnormal.

§ 2.—The Budget and its Compilation.

The Budget, which is awaited so eagerly in these days of heavy taxation, would be impossible of production were it not for the work of the Statistical Sections of the various Departments concerned. These Departments render statements to the Treasury showing the population, who, by a comparison with previous records, are able to estimate the number of people who will be called upon to pay Income Tax, while the average amount of taxable incomes will enable an estimate to be made of the yield of an Income Tax at a certain figure. Figures as to the importation of goods are also available, and hence the yield of any additional Tax upon any commodity imported can easily be ascertained, although allowance has to be made for the fact that consumption may decrease when an increase of price takes place. In the case of Estate Duties the average over a large number of years is the only safe basis to assume, since so many influences may affect the estimate.

Why Results differ from the Estimates.

It may, of course, be argued that the Chancellor of the Exchequer seldom balances his receipts and expenditure, having either a deficit or a surplus. Since this

is so, the science of statistics has not proved of much avail. It must, however, be borne in mind that the Budget is merely a collection of estimates of receipts and expenditure. Receipts may fall from a variety of causes, such as decreased consumption of taxable commodities owing to the increase in the price necessitated by the increased taxation ; and while the average death rate may be maintained, the estates of the deceased persons may be small, and the Exchequer suffer from the smaller yield in respect of Estate, Legacy and Succession Duties. On the other hand, the expenditure of the various Departments may exceed their estimates, or some severe and unforeseen crisis arise which necessitates additional expenditure by some branch of the activities of the Government.

This discrepancy between estimates and results must of necessity always appear, however skilfully the estimate is made; but in estimating on scientific lines, such as those laid down for statistical research, past experience is being drawn upon, and applied to conditions known to be in existence at the time the estimate is made; consequently it is likely to be very much nearer the actual result than a mere guess.

§ 3.—Object of Statistics.

Statistics have been defined as the "Science of Counting."

The ultimate end of Statistical research is to enable comparison to be made between past and present results, with a view to ascertaining the reasons for changes which have taken place, and the effect of such changes on the future. Data relating to past experiences are therefore collected, collated and co-ordinated in such a form as to supply a panoramic view of large

masses of facts easily comprehensible to the man in the street. The experience so ascertained is applied to present conditions, and the current problems solved in the light of such past experience and application. It will thus be seen that while Statistics is a science of estimates and probabilities it is very closely allied with economic progress.

§ 4.—Statistics in Commerce.

So far the science of Statistics has not been directly applied to business except in a few specialised branches, such as Banking, Insurance and Railway traffic management, and even then has been largely confined to the financial aspect of affairs. Business men, however, are constantly drawing on their past experience to solve their problems, though they do not do so through the medium of statistical methods. If the object outlined above is carefully kept in mind there is no doubt that it can be applied equally well to all branches of commercial activity. A commencement has already been made in Cost Accounting, which at its base is the utilisation of recorded information of past work done, and upon such information, allied to a knowledge of present conditions, present or future work may be based.

Business men usually obtain information from the Accounts presented to them from time to time. Modern Accounting may be regarded in a large measure as an exact science, and hence, since the results of a business man's activities are usually presented in the figures compiled by his Accounts Department, there is a tendency to regard business as an exact science also. Accounts, however, record only the financial aspect of transactions rather than the nature of the transactions themselves ; consequently if such figures

only are considered, fallacious conclusions may be drawn.

A moment's reflection will at once make it clear that business itself is composed of estimates and probabilities. A business man buys or produces his goods on the probability of the demand being of such a magnitude that he will be able to dispose of his output or purchases at a profit. The size of his order, or the quantity which he will produce, will depend upon his estimate of the probable size of the demand from his customers. He will willingly undertake to pay a high price for the commodities in which he deals if he is of the opinion that he will be able to obtain a still higher price from his customers. In other words, he estimates the demand, and then buys or produces accordingly.

His estimate may be incorrect. He may overstock himself, and have either to keep unsaleable stock on his shelves or else sell for less than cost. On the other hand he may under-estimate the demand and find that his stock is exhausted before the demand is supplied, thus losing an opportunity of making profit.

A contractor may estimate the cost of the work, and enter into a contract only to find that he has made a loss owing to an error in his calculations as to the cost. On the other hand, his estimate may be too high, and thus he does not obtain the contract for which he has tendered.

The successful business man is the one whose estimate most closely approaches accuracy.

It may at once be said that to estimate correctly calls for skill of a very high order, and this is true, but such skill is only obtained as a result of long experience, and consequently is really the utilisation of past knowledge. This personal skill could be

greatly increased if this past knowledge were carefully preserved, recorded and adequately tabulated, and such recording and tabulating is nothing less than the compilation necessary to obtain Statistics. If records are made over a long period a Standard can be obtained, and such Standard utilised as the basis of future operations; but these operations must in turn be recorded, in order that any change in conditions may be allowed to exercise its influence correctly, and anything abnormal duly eliminated. The compilation of such records would not be an expensive process, seeing that details appertaining to every transaction are available in the office, while records of the financial results are carefully maintained. All that is necessary to establish a satisfactory statistical system is the maintenance of records of the transactions themselves, *e.g.*, the quantities of goods produced, purchased, or sold, etc., and the correct tabulation and utilisation of these records. Care in the compilation of the records is, however, necessary, for unless they are prepared by one who is in close touch with the general policy of the business they may give a biassed view, instead of supplying that broad bird's-eye view of the business as a whole which is so essential to a busy principal.

Large business houses have found it an advantage to establish special statistical and intelligence departments which gather and co-ordinate facts and figures not only of the particular house, but of trade and finance generally. The principals then have access to data which has proved of great value when new markets are being exploited or a new venture launched.

SYNOPSIS TO CHAPTER II.

COMPILATION OF STATISTICS.

§ 1.—The Unit of Data to be collected

2.—Homogeneity of the Data.

3.—Completeness of the Data

4.—Stability of the Data.

5.—Representative Data.

6.—Methods of Sampling.

7.—Methods of Collecting Data.

(*a*) Direct Method.

(*b*) The Questionnaire.

(*c*) Published Data.

8.—Fallacious Comparisons.

9.—Units for Comparison.

10.—Commercial Uses of Statistics.

11.—Divisions of Statistical Data.

12.—Summary.

CHAPTER II.

COMPILATION AND USES OF STATISTICS.

If proper use of Statistical data is to be made it is essential that the collection and recording of data be very carefully carried out, for if this foundation be incorrect the conclusions drawn when data are applied to problems will be fallacious, and the action taken give results very greatly at variance with the estimate made. The collection of data therefore may be said to be the foundation of Statistical Science.

Before collecting data of any description it is very necessary to consider the object for which they are required—and then all possible information bearing on such object is gathered, collated and tabulated. To collect data without an object in view is, generally speaking, a waste of time, for it will lack that co-ordination so necessary to successful application. It has been stated that "Figures can be made to prove anything," but this is due either to wrong methods of collection, incomplete data, or in data not being comparable. Figures themselves cannot prove anything. It is the method of presenting and using them which results in correct or incorrect conclusions being drawn.

§ 1.—The Unit of Data to be Collected.

It is a truism to say that definitions vary materially. Take the term "profits" as an example. What are profits ? If this question be addressed to a business man his answer will vary greatly from that

obtained in reply to the query from, say, an Accountant; while an Inspector of Taxes or a Lawyer would each give a rendering entirely different from either of the others. The profits of a business, as calculated by a Vendor of such business, will frequently materially differ from those calculated from the same data by the person negotiating for its purchase. Each would be correct, according to the definition he had taken, but to compare figures obtained from each of these various sources would lead to no intelligible result for statistical or practical purposes.

From a Statistical point of view comparability will be possible when all the data conform to the same definition, no matter which definition be taken. In view of conflicting views on some points it will probably be better to take profits as calculated for Income Tax purposes, since Inspectors of Taxes follow certain hard and fast rules, thus tending to uniformity throughout the country.

In statistical work, therefore, it is first necessary to fix upon a unit such as it is desired to record, carefully define it, and then proceed to collect the data, seeing that the definition is strictly adhered to in every case. Even though the definition be inaccurate, the data are comparable if they are homogeneous, and hence their use will not result in fallacious conclusions being drawn.

§ 2.—Homogeneity of the Data.

Homogeneity may be said to be the first requirement of collection to be complied with, and it follows therefore that unless the unit chosen is applicable to all the cases which fall under review, it will be necessary to split the unit into groups or classes in order to get more accurate results.

Wages form a good example of such a unit. In its broad acceptation in the business world the word has come to mean the remuneration paid to those workers engaged in the actual handling of the commodities in which a business engages, and thus the pay of clerks, etc., falls into the category of salaries. This seems so clear that it might be thought sufficient, but in practice it will be found that unless all the workers are paid upon the same basis comparison is rendered difficult, if not impossible. Where different rates are paid, the workers should be grouped according to the basis of pay.

As an example let us consider a case where both male and female labour are employed, and assume, as is usually the case, that the rate for women is lower than that for men. If any change in the composition of the operating force takes place, such as frequently happened during the Great War, when male labour was reduced and female labour greatly increased, the average wages paid to each worker would, all other things being equal, fall, unless the wages were split into the two groups: "Wages paid to Men," and "Wages paid to Women," in which case any change in the relative numbers employed will not change the average.

Take a concrete case. A firm employs 100 men and 500 women, and pays to the former a total of £500, and to the latter £1,000 per week, *then* the average

wage per man is $\frac{500}{100}$ or £5 per week,

and per woman $\frac{1,000}{500}$ or £2 per week.

If the pay roll is combined, there are 600 employees earning £1,500, an average of £2·5 per

week, an average which does not represent the true state of affairs. Now, suppose the number of men employed by the firm *falls* to 50, with a weekly pay roll of £300, and the number of women becomes 1,000, with a wage bill of £2,325. The average earnings of the men is now £6 per week, and of the women £2·325 per week, an increase in the average earnings in each case. If, however, separate records have not been kept, the wage bill would be £2,625 for 1,050 persons, or an average of £2·5. If this latter fact were all the information at our disposal, we should assume that no change in the rates of pay had taken place, whereas, eliminating any question of overtime, the rates had increased considerably.

Similarly it is obviously fallacious to compare the profits of two periods if the amount of Working Capital in the two periods is not the same, nor would any end be served by comparing the profits of two periods if in one of them some extraordinary profit had been realised, such as, say, from the realisation of some investments. If the best results are to be obtained, the profits must be arrived at after the compilation of the accounts on the same basis, *i.e.*, after crediting the usual trading profits, and charging the usual items, or at any rate the same class of item.

Homogeneity is absolutely essential in any tables, not only in the returns at any one time, but in the returns for the whole of the periods comprised in any one table. If it should become necessary, as is sometimes advisable, to change the basis of compilation, care must be taken to note the change, as otherwise fallacious observations would be made. Too many notes, however, are not advisable, for they render the table cumbersome to use, and greatly impair its utility. An example of the change of

base is to be found in the Tables of "Quantities of Wheat, Wheatmeal and Flour, etc., imported into the United Kingdom," published in the "Statistical Abstract of the United Kingdom." In 1904 and subsequent years the figures were classified according to the "Countries of Consignment," instead of under the "Countries of Shipment," as in the earlier years, a most important alteration which particularly affects those figures relating to countries which either have no ports of their own, or no regular direct shipping connections with this country, and also those countries which act as intermediaries in shipments from the first named. These latter countries will show considerable reductions in the supplies of such commodities to this country, while the former will be greatly increased.

If reference is made to the Tables in the Abstract it will be seen that the exports of such commodities from the United States apparently fell from 24 million cwts. of Wheat and 16 million cwts. of Wheatmeal and Flour, in 1903, to only 7 million, and 9 million cwts. respectively, in 1904. On the other hand, the Argentine Republic was only given credit for supplying us with 14 million cwts. of all kinds of Wheaten commodities, in 1903, but in 1904 the amount had risen to 22 million cwts. If no note were appended to the Table, the figures as they stood would have caused an observer to come to the conclusion that America's production of that class of commodity was falling, and that as a consequence we were compelled to resort to the Argentine to make good the deficiency; whereas it is more than probable that the output in both countries was normal, and that the variation shown was entirely due to the alteration in the basis of compilation.

§ 3.—Completeness of the Data.

It is also essential that the data collected be complete in respect of all matters appertaining to the unit which need to be also collated and co-ordinated. If this be not done we shall obtain inaccurate inferences. If we are collecting data relating to the profits of companies engaged in certain industries, we must also collect information regarding the Capital employed in such companies, for, *all other things being equal*, the company with the largest capital should make the largest amount of profit. Thus, if we were told that the profits of the A. Co. for a particular year were £20,000, while those of the B. Co. amounted to £100,000 the general assumption would be that B. Co. had done five times as well as the A. Co. If, however, the capital of the A. Co. is £100,000 while that of B. Co. is £1,000,000 the conclusions will be quite different for A. earns 20 per cent. on its capital against B's. 10 per cent. In connection with the capital we are faced with the difficulty of defining this, as already mentioned in § (1) above, particularly in the case of Limited Companies—seeing that in Company matters there are various forms of Capital, such as Authorised or Nominal; Issued, Subscribed, Paid-up, Reserve, and Working Capital. In order to ensure complete uniformity, however, it is always wise for the purposes of comparison of any kind to take as the Capital *the actual interest of the proprietors.* This can be arrived at by deducting from the total value of the assets the liabilities of the business to all third parties. In the case of a Limited Company the interest of the proprietors will usually be composed of three items, viz., (1) Issued and Paid-up Capital; (2) Reserves; and (3) any balance of Profit and Loss. Reserves are accumulated from undivided profits and so increase the

interest of the proprietors and the actual capital of the concern. Similar remarks apply to any credit balances on the Profit and Loss Account. If this last named Account has a debit balance the amount must, of course, be deducted from the other items. Even this method of arriving at the capital is hardly satisfactory for statistical purposes, for the valuations of the assets may be on different bases and hence the capital be affected. Uniformity in methods of valuation and in classification of items would greatly aid the statistician dealing with commercial data.

If certain factors only are compared, errors are obviously likely to arise, but if the science of Statistics is carefully studied these errors are not likely to be made. All the factors must be studied if the real facts of the situation are to be deduced. Every unit must be considered and given its due weight, for if this be not done we shall be arguing that since one part is progressing the whole must be well. Examples of this kind of reasoning can easily be called to mind. For instance, during the early part of the year 1920 the value of our exports showed increases, and many of our newspapers forecast a more prosperous time. They, however, omitted to take into consideration the ever increasing prices, and until tables of quantities were available as well as the monetary value of our exports we were unable to say whether the volume of trade had increased or not. It has also been argued that because our Export Trade is improving, the country must be in a flourishing condition, but this is not necessarily so. The Export Trade is only a part of the business of the country, and frequently when the Home Trade is suffering from depression the Export Trade shows an actual increase, owing to the fact that manufacturers and merchants always endeavour to obtain the best price for their commodities, and if the

Foreign Markets will yield a better return than is obtainable at Home, then the goods are shipped to other countries.

Another example is frequently met with, especially at the present time of high rates. The rates per £ are compared without any cognisance of the rateable values being taken. It frequently happens that high rates in one borough are accompanied by low assessments, while in another the rates may be low and the assessment high. If we take two similar houses in two boroughs, where the rates are respectively 15/- and 12/- in the £, it would be argued that the second is better for the ratepayer; but if the assessment in the second is 25% higher than in the first named the actual amount of rates paid will be the same in each case notwithstanding the disparity in the rates levied.

It is not only necessary to record other data directly relative to the unit under review, but all factors which are likely to have even the slightest effect upon the results must be watched for and noted. Bank Holidays, for instance, since they affect output, need to be recorded, for to compare the output of a factory for two weeks, one of which contained a Bank Holiday, would lead to entirely erroneous conclusions were the fact not known. Similarly it is not sufficient merely to know how many men are employed in two different factories; we must also know the average number of hours worked per day or week, if a comparison of output is to be made. The methods of production used in the factories would also need consideration, for two factories producing the same kind of article, and employing the same number of hands for the same period of time, may show entirely different outputs, owing to one having better methods of production than the other. Other examples are given later.

§ 4.—Stability of Data.

As will be seen later in this book, observations over a short period, while useful for investigating reasons for special movements in the data, are unsatisfactory for use as a Standard for comparison, as they are apt to be seriously affected one way or the other if abnormal conditions are present. By collecting data over a long period of time the effect of abnormalities in one direction is apt to be offset by abnormalities in the opposite direction, hence a standard suitable for use in comparisons can be obtained. It follows therefore that any unit taken must remain constant in definition or value during the whole of the period being investigated, as if any change takes place the results are bound to be affected. Absolute stability of the unit is practically impossible of attainment, but this can to some extent be compensated for by the use of the Moving Average described on page 77 *et seq.* Capital is a unit which has but little stability if the definition on page 16 be considered, seeing that most Companies find it expedient to increase reserves from time to time, while the balance of profit and loss carried forward varies materially, and thus the real Capital of the concern alters with each accounting period.

§ 5.—Representative Data.

In any extensive research undertaken it is obviously impossible to investigate every example in existence, so it becomes necessary to obtain a selection and investigate these, without going to the trouble and expense entailed in dealing with the whole of the individual cases. This method of procedure has been adopted in social and physical investigations and is based on the "THEORY OF PROBABILITIES," which

states that if a moderately large number of items be *chosen at random* from among a very large mass, such numbers are, *on the average*, almost sure to have the same characteristics of the larger group, and the data so obtained can be safely used as a base for comparison with all other examples of the same kind. Some statisticians call this principle the "LAW OF STATISTICAL REGULARITY," and it is largely used in Insurance. Similarly, data obtained over a long period are more reliable than those collected over short periods, as abnormalities in one direction are offset by those in the other and opposite. This is known as the "LAW OF INERTIA IN LARGE NUMBERS" and is dealt with in detail later. The main point to be borne in mind is that the examples taken must be representative of the whole body which it is desired to investigate; that is to say, every unit of the mass must have the chance of being included in those which are actually investigated. If, for instance, it is desired to ascertain the average rate of profit earned in any particular trade, and the number of businesses in such trade are so numerous that it would be impossible, too laborious, take too much time, or prove too expensive to apply to all of them for the requisite information to enable such average to be compiled, then by circularising a sufficiently large number of them to ensure a reliable average being arrived at, a reasonable estimate of such average rate applicable throughout the trade could be made. Every business must have a chance of being included, whether it be large or small, and in order to ensure this being so, perhaps the best method of selecting the representative businesses would be to draw them as is done in a lottery. This process of selection is known as "Sampling." If only large businesses were included in the inquiry the average would be fallacious,

in so far that a large business enjoys many advantages over a small one, such as cheaper prices when buying in large quantities, or a lower ratio of expenses to Turnover. Some businesses may get more favourable terms against payment in cash, whereas other houses may not have sufficient working capital to permit of large cash purchases, and thus suffer some disadvantage. We should not measure the stature of the British Army by the average height of the men in a Guards Regiment, neither do we measure business results by taking the large businesses only. In a general inquiry of the nature outlined above care must be taken to see that every type of district is also represented, or has the chance of representation, for conditions vary greatly, and the rate of gross profit earned in one part of the country may be exceeded in another, or reduced in a third. Businesses which trade for cash only, as well as those which deal on credit, must be included, for the rate of gross profit in the latter will, of necessity, need to be higher than in the former, in order to cover the risk of Bad Debts.

§ 6.—Methods of Sampling.

By using the Lottery method for the selection of cases to represent the whole of a trade or group, the Law of Probabilities will give all units, examples, classes and districts the same chance of inclusion, though other methods of sampling may be equally efficacious, such as taking every tenth unit or case in the list, or pricking the requisite number of examples from the list while blindfolded; but a method of personal selection of the representative cases is not satisfactory, for unconsciously there may be some bias present, which leads to the inclusion of some examples and the omission of others.

Naturally the larger the number of representative examples taken and investigated the more satisfactory

the standard becomes. The success of Life Insurance is very largly due to the very long period of time over which the statistics have been compiled, and the variety of trades and professions followed by the policy-holders. At present the science of Statistics is only just being appreciated as an aid to business, for until the last few years it had been chiefly applied to social problems and branches of business, such as Insurance, arising out of social conditions. It must, however. be obvious that the ordinary business does not offer so wide a scope for the collection of data as is advisable, or as that which is possible in social statistics. On the other hand, the smaller data available will, or should be, more accurately compiled than those which cover a wider area, and consequently can be relied upon to a much greater extent. Care, however must be taken in such cases not to argue from the particular to the general, as the facts applicable to one business may not, and often do not, apply to others of the same kind elsewhere.

§ 7.—Methods of Collecting Data.

In collecting the necessary data to enable the purpose of the collection to be achieved, the first essential is to decide what information is required, and then define the unit it is desired to examine and compare, care being taken to see that it conforms to the rules already laid down. The field of investigation must next be decided upon ; the exact scope of the investigation defined ; and the process of collecting the data undertaken. It is essential, as has been previously pointed out, that in order to obtain useful and reliable data the unit in all the cases under inspection must be arrived at in the same manner, so that if the enquiry is in regard to profits, care must be taken to see the

profits are in all cases calculated in the same manner, otherwise it will be found that some people will return Gross Profit, others the Net Profit, while some will bring extraordinary profits or losses into account, so materially affecting the results as to render comparison ineffective. The profits adjusted for Income Tax purposes will be found to have been returned in some cases, and divisible profits may appear in the returns of companies. In some cases Depreciation will have been allowed for, while in others it will not have been charged. Allocations to Reserve may have been made in some examples, while in others no such allocation may have been made. It follows therefore that not only must the unit be carefully defined, but the methods to be adopted in obtaining data conforming to that definition must be laid down.

(a) Direct Method.

Statistical data, should, whenever possible, be gathered by the compiler personally, or by those trained under and supervised by him, thus ensuring uniformity in the data, and in the methods adopted for collecting, sampling and tabulating. This course is seldom possible owing to the enormous amount of time and/or expense involved, and recourse has then to be made to the circularisation of Schedules of Questions (technically known as "Questionnaires").

RULES FOR COMPILING QUESTIONNAIRES.

(1) They must be very carefully compiled and the unit explained concisely but clearly, if mistakes, misapprehensions and misunderstandings are to be avoided.

(2) Whenever results involving calculations are required it is always advisable to ask for the original data, so that the compiler may make

the necessary calculations himself, thus reducing the chances of errors in obtaining the result, and ensuring uniformity in the method of calculation adopted.

(3) The questions in the schedule must be short and clear, easy to understand and answer, and few in number.

(4) On no account should the questions call for information of a confidential nature, as otherwise the whole of the questionnaire will probably remain unanswered.

(5) Care must also be taken to see that the questions can raise no resentment in the mind of the average man, so they should not be of an inquisitorial nature.

(6) They must be free from ambiguity and correctly and clearly define the information desired.

(7) If an opinion is required the question must be so drafted that the answer to be given should be "Yes" or "No" without qualification.

(8) Questions should wherever possible call for answers of a corroboratory nature.

If these points do not receive attention many of the schedules will not be returned by the people to whom they are addressed.

(b) The Questionnaire.

Assuming that it was desired to ascertain some idea as to the effect of increased Wages upon output in a given industry, the following might well be the Questionnaire circulated, assuming that the articles produced were of uniform standard and size, or that the same process, where several processes are necessary to give a complete product, is referred to only. Since the form is a general one it would need adaptation to a particular business.

Name of Firm :

Place of Business :

Description of Article Produced :

Method of Production, or, if various processes are necessary, the process referred to :

Has any alteration been made in the Method of Production, or any improvements in the Plant effected, with a view to speeding up Production? If so, give details, date of coming into operation, and the effect upon Output :

Has there been any breakdown of machinery, or other extraordinary cause of delay, which would affect Output? If so, state the nature of such delay, the number of hours during which Output was affected, and the extent of decrease :

Date of Increase of Wages :

Basis of Increase :

Basis of Payment of Overtime :

Method of Payment—Time or Piece :

Has any alteration been made in the Method of Payment during the periods under review? If so, give details :

	For Period of Six Months prior to Change	For Period of Six Months subsequent to Change.
1. Number of Men employed :		
2. Total Hours worked :		
3. Total Wages paid :		
4. Number of Hours of Overtime worked :		
5. Wages paid in respect of such Overtime (included in question No. 3).		
6. Total Output :		

(c) Published Data.

Frequently direct enquiry is out of the question, and recourse has therefore to be made to such published statistics and data as are available. This does not usually give such satisfactory results to the observer as the direct or personal method of compilation, since the methods used in the compilation are unknown, and there is usually nothing to demonstrate the homogeneity of the information. Wherever it is found that figures are unreliable they should either be discarded altogether, or adjusted, if adjustment be possible. It must always be borne in mind that a smaller number of correct figures will enable a more accurate estimate to be made than a larger number of incorrect ones.

The Board of Trade as well as other Government Departments supply much valuable material in their various returns, but as a general rule the business man either has not the time or the inclination to make a study of Blue Books. It must also be admitted that many of these figures are compiled in such a manner that they are not so useful to the business man as they might, or as they should be. This has long been recognised, and some years ago a special conference of Government Statistical Officers of the British Empire was held, under the auspices of the Board of Trade, to consider what steps could be taken to ensure homogeneity in the returns, and uniformity in the method of compilation. Among their recommendations was one relating to the formation of an "Empire Statistical Bureau," which should be empowered to demand statistical information of public importance, and should compile and publish from time to time monographs embodying important trade statistics, such as those dealing with: Trade,

Production, Transport, Communication, Population. Labour and Industries, Finance, etc. Such a Bureau, if and when formed, should prove of great assistance to the business man, but its usefulness will be greatly impaired unless general interest is taken in the subject by the business world as a whole, since unless the statistics compiled are used for business purposes their compilation will be a useless expense and waste of time. The business man needs to study this question in the broadest possible sense, and without prejudice. He should study the returns, and see what information they contain which either in their present form, or with the aid of additional calculation, can be used to assist in placing his business upon a scientific footing. Such statistics might also open up the possibility of new fields for the employment of capital and labour, as for instance, when the returns of imports show a large figure for some commodity which might, with the aid of skilful organisation and adequate financial resources, be quite as easily produced in this country. This possibility was forced upon us during the Great European War, when industries sprang up here to produce goods previously imported from enemy countries.

Statistical information which would be of great use to the business man is rendered difficult of compilation by the reluctance of individuals or firms to disclose information which they think will be of use to a trade competitor. This reluctance is, in a way, natural, but the benefit to be derived from properly prepared statistics would more than outweigh the disadvantages attendant upon the disclosure of important data. If the great utility of Statistics as an aid in scientific management were only better recognised, the formation

of a Statistical Bureau, under the supervision of a Chamber of Trade, or Commerce, where data appertaining to any particular industry, and trade generally would be collected and collated for the benefit of the business world, should prove of inestimable advantage. The Federation of British Industries has already established such a Bureau, which deserves greater support from the members.

§ 8.—**Fallacious Comparisons.**

It is care in the use of the data at our disposal that will make a Statistical Department useful, and avoid the pitfalls of fallacious comparisons, of which some examples have already been given in this chapter. Many business men compare the results of their year's trading with the accounts of the previous trading period, but such comparisons are to a very large extent useless, for either of the two periods under review may have been affected by abnormal conditions and thus the perspective be thrown out. Every trade has its cycles of good and bad trade, and if it should happen that a very good year is followed by a period of acute depression, fallacious conclusions would be drawn if the trading results of such two periods were compared without taking into consideration the general state of trade. Enquiry, however, should be made with a view to discovering whether the fall in the results is solely due to adverse conditions, or whether it is due in whole or part to internal causes. Comparison with the published results of companies in the same trade will form a useful guide on this point. Similarly when a good year follows a poor one, optimism must be restrained, since the greatly enhanced results may not be maintained. It is the welfare of the business over a long period which needs to be studied, and if this is

progressing then the rises and falls which are bound to occur are relatively of but little importance.

Again, the monetary value of the Turnover of a business may show a considerable decrease, but this does not of itself prove that the business is on the down grade. Prices may have fallen abnormally, and the actual turnover of goods be the same, or greater, than in the previous period. In the case of the retail trade the selling price of commodities is usually fixed at Cost plus a percentage, such percentage being calculated to cover expenses of conducting the business and leave a margin for profit. While it is true that when competition is keen prices have to be lowered (but so long as the price does not fall below a figure sufficient to cover the original cost of the goods and the percentage to cover the expenses of management, and the quantity sold produces such costs and expenses, no loss is incurred), yet the original Selling Price is fixed in the manner indicated. The effect of a reduction in the general level of prices may be easily demonstrated if we assume that in two trading periods the same *quantities and qualities* of goods were bought and sold, but that in the second of the two periods the general level of prices fell by 20% (*i.e.*, Commodities taken as a whole fell 20% in price). If the Selling Price be Cost Price plus 50%, then the following statement gives the results in Comparative form:

	1st Period.	2nd Period.
Cost Price of Goods	£120,000	£96,000
Selling Price of Goods	180,000	144,000
Gross Profit	60,000	48,000
Expenses	45,000	36,000
Net Profit	£15,000	£12,000

It has been assumed that the expenses have also fallen by 20%, but this is an unlikely contingency—seeing that economic research clearly proves that expenses move less violently and at a considerably later date than the general level of prices; so that if cognisance were taken of this fact the result in the second period would be even worse than that shown above. Even so it will be seen that the results show a serious diminution in every way, and yet the *actual volume of* business transacted was the same. The economist will of course point out that even though profit does show a decline of 20%, yet, as the general level of prices has declined to a like degree, the purchasing power of the profits in the second period tends to be the same as those of the first. While this is quite true, the average business man considers results as shown in the terms of money rather than in the purchasing power of that money.

It will at once be recognised that in such a case as the above the business might have sold 25% more goods in the second period and only show the same monetary turnover as in the first. This state of affairs is one of the reasons why the published results of many companies for the past few years show a great diminution of profits, if not actual losses.

It will thus be seen that at the present time it is impossible to compare the figures of Turnover or Profit shown for the current period with any of the War years, unless at the same time we know exactly what change has taken place in the general level of prices, or unless we are acquainted with the quantity of goods that has been sold in each of the periods under review.

Similarly, if the profit be lower than in a preceding period, though the turnover of goods was the same,

it would not be correct to assume, without full enquiry, that prices were lower than in the previous period. They may have been increased, but the working expenses increased out of all proportion to the increase in turnover, leading to a fall of profits. Many other influences may affect any or all of the factors.

It will be apparent that in order to avoid fallacious conclusions we must collect all data which bear even remotely upon the unit selected for investigation, as these, as shown above, may have a very marked effect.

During the past few years prices have fallen very considerably, and many businesses are showing diminished profits, if not actual losses, as a result. The example given above is therefore by no means an impossible or improbable one, as is illustrated by the following extract from the speech of Mr. H. Gordon Selfridge to the Shareholders of Messrs. Selfridge & Co., Ltd., at their annual meeting to consider the accounts for the year 1921.

> "1921, for the most part, *has been a year of declining prices* and shows a growing disinclination on the part of the public to buy any but the real necessities.
>
> "Our increase in yardage or in the number of articles sold over 1920 or any previous year is very large. The actual increase in customers or in individual transactions is over one million seven hundred thousand, *but the decrease in values has made it less easy to show an increase in actual returns.*"

Occasionally we find estimates of the profit earning capacity of a Company based upon the output of its factories or works, but profits do not depend upon output, but upon the quantity of such goods sold and the price which is obtained for them.

§ 9.—Units for Comparison.

It has already been stated that it is invidious to compare the output of a factory for a week in which a

Bank Holiday occurs with that for a week in which full time is worked. Similarily it is impossible to compare the figures of output for two weeks when in one of them more labour is employed, or longer hours worked, than in the other. How then can we compare output if conditions vary from time to time, seeing that homogeneity is impossible? It is necessary in such cases to prepare a unit which while in itself constant includes the variable data upon which the output depends. In the case of output it is necessary to combine both time and labour. Such a unit would be the "Man-Hour." The number of Man-Hours during which output was in progress is the number of workers multiplied by the average number of hours worked per individual, or in other words the total time occupied by all the workers. Thus one man working 30 hours would give us 30 Man-Hours, while five workers each working on the average six hours would also give us 30 Man-Hours. It will be at once apparent that we have now a unit which is applicable to the business employing few as well as many hands.* To obtain the output per Man-Hour all that is necessary is to divide the total output by the total Man-Hours or total time worked $\left(i.e., \frac{\text{Total Output.}}{\text{Total Man-Hours.}}\right)$. When this is applied to two or more factories the resultant of such computation is a comparable factor. The data resulting from the circulation of the questionnaire on page 25 would be best treated in this way, the output for each plant per man-hour being calculated for the periods before and after the change, and comparison of the figures thus obtained would give us the desired

*It must not be overlooked that the Man-Hour is not necessarily applicable in all cases, as, for example, many industries necessitate the employment of gangs of workers to do particular jobs. In such cases the unit would have to be "Gang-Hours." The principle is, however, the same as for the Man-Hour.

result. Similarly if the total wages paid were divided by the Man-Hours worked we shall get the average wages paid per Man-Hour, and comparison of the rates of wages paid in relation to the output per Man-Hour will be rendered easy. If, however, overtime had been worked at a remuneration in excess of that ordinarily payable, a deduction in respect of the amount paid in excess of the normal rates must be made from the total wages paid if the figures are to be homogeneous and represent the standard rate of pay. Naturally where the method of production varies separate Tables would have to be prepared for each method or process of production.

Such a standard as the Man-Hour will eliminate the consequence of accidents resulting in a total or partial suspension of work, as the time lost will be deducted from the total time, and the breakdown will have no effect upon the output per Man-Hour, though, of course, the total output will of necessity be reduced.

If the maximum information is to be obtained, the figures appertaining to the Total Output and Output per Man-Hour would be best tabulated in five columns, the first showing the number of men employed; the second the average number of hours worked per employee; the third the number of Man-Hours (this is obtained by multiplying together the figures in columns 1 and 2); the fourth the total output; the fifth the output per Man-Hour, obtained by dividing the figures in Column 4 by those in Column 3. The data for each week or other period is then capable of very easy comparison. If the figures in column 4 fall, while those in Column 5 remain constant, attention will immediately need to be given to the reason therefor. The fall in the total output may be due to one or two causes, or to a combination of

both ; viz., a fall in the number of employees, or a decrease in the number of hours worked, but this can be instantly traced by reference to Columns 1 and 2 of the form. Should the factors shown in the first two columns remain constant, but both the Total Output and the Output per Man-Hour show diminution, inquiry must be directed to the cause, which may be due to inferior and inefficient workers having taken the place of better ones, or to physical or mechanical difficulties being present and affecting output.

The following example will serve to illustrate the utility of such a statement :—

	(1)	(2)	(3)	(4)	(5)
Week ended	No. of Employees.	Average No. of hours worked per Employee.	No. of Man-Hours.	Total Output in Tons.	Output per Man-Hour in Tons.
January 7th ..	200	45	9,000	4,500	·5
,, 14th ..	190	44·7	8,493	4,077	·48
,, 21st ..	200	44	8,800	4,400	·5
,, 28th ..	200	45	9,000	4,545	·505
February 4th ..	250	44·8	11,200	4,480	·4
,, 11th ..	250	45	11,250	5,062·5	·45
,, 18th ..	260	45	11,700	4,914	·42
,, 25th ..	260	45	11,700	5,382	·46
March 3rd ..	255	45	11,475	5,508	·48
,, 10th ..	260	40	10,400	5,200	·5
,, 17th ..	260	45	11,700	5,967	·51
,, 24th ..	260	45	11,700	6,084	·52
,, 31st ..	260	44	11,440	5,948·8	·52
Total of Man-Hours for 13 weeks			137,858		
Total Output for 13 weeks ..				66,068·3	
Output per Man-Hour over 13 weeks					·4792

By adding the figures in Column 3 we obtain the total of the Man-Hours in the period under review,

while the total of Column 4 gives us the total output for the same period. By dividing the total Output (*i.e.*, 66,068·3 tons as shown in Column 4) by the total Man-Hours (137,858) we obtain the output per Man-Hour for the whole of the thirteen weeks shown. This figure (·4792 tons) would then form the basis of comparison for succeeding weeks. At stated periods this output per Man-Hour should be adjusted by including the cumulative Outputs and number of Man-Hours to the new date, and obtaining a figure for the Output per Man-Hour for the whole period. Care must be taken to calculate the correct Output per Man-Hour, as it is not correct (for reasons to be shown later), to average the standard results obtained separately for the periods. Supposing that in the succeeding 13 weeks the Total Output was 84,000 tons, and the number of Man-Hours was 140,000, then the Output per Man-Hour for that period will be ·6 tons. To obtain the Output per Man-Hour for the whole of the twenty-six weeks we must add the Total Outputs and Man-Hours for the two periods, and find the Output per Man-Hour from these figures, *i.e.*:—

	No. of Man-Hours.	Total Output in Tons.
1st Period	137,858	66,068·3
2nd Period	140,000	84,000
Totals for 26 weeks	277,858	150,068·3

then Output per Man-Hour for the 26 weeks is

$$\frac{150{,}068\cdot3}{277{,}858} = \underline{\underline{\cdot54009 \text{ tons}}}$$

whereas, if we average the Outputs per Man-Hour we get:—

$$\frac{\cdot4792 + \cdot6}{2} = \frac{1\cdot0792}{2} = \underline{\underline{\cdot5396}} \text{ tons.}$$

This result is entirely incorrect, as can be easily demonstrated by multiplying such result by the number of Man-Hours. This should give the Total Output for the period, but instead of doing so gives only 149,932·1768 tons, whereas by multiplying the result obtained in the first answer by the number of Man-Hours we get the actual total output.

By so adjusting the results the unit for comparison is amended to include current working, and a more reliable and valuable standard is obtained, in so far that any abnormal influences are spread over a longer period, and a poor output in any week is counter-balanced by one the production of which is above the general level. This is very necessary, particularly where conditions change, as experience shows that work-people cannot be expected to maintain a constant and steady output under varying conditions. If the number of work-people vary greatly, or the average number of hours worked per employee fluctuate to any great extent, the Output per Man-Hour should be adjusted week by week, as otherwise the unit for one period would not be the best standard for comparison. By constantly adjusting this standard any improved methods of production, or greater efficiency of the workers resulting in a higher output will be given due weight.

In the statement given above it will be noticed that a very heavy fall in the production per Man-Hour took place in the week ended 4th February, and that subsequently, with one exception (*i.e.*, the week ended 18th February), the output per Man-Hour steadily improved until the last week. The abnormally small output per Man-Hour for the week ended 4th February may be due to one or both of two factors, which are revealed in the statement. The first is that

during the week in question the number of operatives increased from 200 to 250, and it is not only possible, but highly probable that the new work-people were not able to produce the normal output owing to strange working conditions. In the second place it will be seen that the average number of hours worked per operative had dropped from 45 hours, which would appear to be the normal working week, to 44·8. This reduction in hours may have been due to the temporary absence of a "Key" worker or workers, and production handicapped accordingly. It is, however, noticeable that when the working force was again increased, during the week ended 18th February, the same phenomenon is observable, though in this case the full number of hours per operative was worked, thus confirming the fact that Output is adversely affected when operatives are working under strange conditions. This fact is also emphasized by the Output per Man-Hour steadily increasing during the ensuing five weeks, indicating steadily improving efficiency on the part of the employees. It will also be observed that although the average number of hours worked per employee for the week ended 10th March, showed a heavy fall, yet the output per Man-Hour was not affected, although the Total Output for the week necessarily showed considerable diminution. This clearly shows that such a unit as the Man-Hour is one which can be used to make data, which in itself is not homogeneous, capable of being used for the purposes of comparison.

The utility of this table would be greatly increased were it possible to place at the head of the fifth column the output per Man-Hour for the previous year or quarter, or whatever unit of time is adopted, thus permitting an easy comparison with previously observed

results. While in the science of Statistics it may be broadly stated that a Standard based on a long series of observations is far more reliable than that based on a short term, it is doubtful if this would apply with equal force to a productive business, since working conditions are frequently altered and improved, and output affected accordingly; but if the standard be based upon the whole of the results obtained while the conditions remain unchanged, the standard will be far more reliable than one obtained over a short period of, say, a year, since the longer period will allow for and correct all extraordinary fluctuations. This is the "Law of Inertia of Large Numbers," to which reference has already been made. It arises from the fact that in a large mass of data phenomena are bound to occur, but can be ignored in so far that while one phenomenon will affect the result in one way, another will affect the data in the opposite direction, and consequently tend to nullify the effect of the first, or entirely neutralise it. The larger the mass of figures the greater opportunity there will be for this law to operate, and hence the necessity for basing the standard as far as possible on figures obtained over a long period.

While the form set out above is representative, it will be found that each business presents its own problems, and consequently the business statistician will have to study his own business carefully before he will be able to fix upon a standard satisfactory for all purposes; but once that standard has been decided upon, it should not be changed without very careful consideration, for if it be altered the previous compilation of data will need to be "scrapped" or much time and labour expended in adapting it to the new conditions.

Other units of the same nature as the Man-Hour can easily be constructed. For comparing transport costs we use the *Cost per Ton-Mile, i.e.*, the cost of carrying one ton for one mile; the *Cost per Wagon Mile, i.e.*, moving one wagon for one mile; *Cost per Car-Mile; Cost per Passenger-Mile*, etc. In Cost Accounting such units are essential if the results are to be strictly comparable, though frequently several units will be used in order to compare the results of the production. In the electrical industry for instance it is usual to find:

1. The cost per ton of Coal burned.
2. The cost per unit of electricity generated.
3. The cost per unit of electricity sold.

Comparison of the first two shows the type of fuel which is the best for economical generation, while a comparison of the last two, after allowing for units used on the works will show the loss in distribution which is taking place. In coal-mining we should compare the costs per ton of coal raised, and the costs per ton sold.

If data of this nature were available, sub-divided where necessary by districts, the individual producer could, after compiling data from his own business on similar lines, compare the results of his business with the general results shown, and also by the figures for his own district. From such comparison he could easily see whether he was producing efficiently, and as well as the general result, and if not, make enquiries as to the efficiency of his plant, labour and management, with a view to bringing his own business into line with the general tendency. By such means the possibilities of obtaining greater profits would be pointed out.

A business which maintains a record of the quantities of goods purchased and sold over a particular period

can judge whether the business is progressing or not by studying the figures of such quantities, and thus can base its future production or purchases upon its past experience—obtaining a valuable clue to the general demand of its customers and the periods of time in which such demand arises. This eliminates guess work and leads to stocks being ready to meet demands. By maintaining quantitative records we can gauge the position of a business far more accurately than will be possible when monetary records only are kept. Prices are bound to fluctuate from time to time, or, to make the same statement more scientifically, the purchasing power of money changes, and the quantity of goods handled will act as a check upon the progress or otherwise of the business. Financial Records are essential to a business, but these alone are apt to give rise to very fallacious ideas, especially at the present time when prices show constant fluctuations.

§ 10.—Commercial Uses of Statistics.

Since Supply and Demand fluctuate widely, statistical data should enable a business man to discover when supply exceeds demand. When this happens prices would tend to fall, and knowing this he could postpone his purchases whenever possible till such fall takes place. Similarly when demand exceeds the supply, prices have a tendency to rise, thus indicating the time for judicious selling. His records should supply him with information as to when trade falls off, and so indicate when advertising or special offers are called for to keep business moving. He will discover when popular taste changes and thus not overstock himself with those commodities the demand for which is steadily declining, but will purchase those for which the public is asking. He will notice the

seasonal changes in demand, and adjust purchases and stock accordingly. A business man who is working on borrowed money will, by compiling statistics as to the rates of interest ruling over various periods, discover from his records that at certain times of the year money is always cheaper than at others, therefore if he can plan his operations so that his borrowings are heaviest when money is cheapest he should reap additional profit from the procedure. In normal times foreign exchanges move favourably to this country at certain fairly well defined periods, so that a business man knowing this might be able to arrange his imports, in order that the payments therefor are made at the time when Bills on the exporting country are at a discount. At the present time conditions are not normal and hence past experience is not helpful. Records can be compiled of delivery charges so that the cheapest and speediest method may be adopted, and fluctuations in running costs traced and corrected. The output of machines can be compiled with a view to obtaining a "Standard of Output" to act for all machines working under similar conditions. Variations from this Standard would then be disclosed, and could be enquired into with a view to adjustment, thus leading to greater efficiency from both plant and workers. Tabulated results of Advertising will show the best type of advertisement, the best medium or media in which to advertise, the district or town from which the best results are obtained, and even the best days on which to advertise.

There is no limit to the variety of purposes to which statistics can be put, but it is impossible here to do more than indicate what can be done. The data to be collected, the object of its compilation, and the uses to which it can be put are a matter for the head of each business.

§ 11.—Divisions of Statistical Data.

Statistical Data fall into two main divisions, viz., (*a*) Groups and (*b*) Classes.

(a) Groups.

When a large number of examples of any kind have something in common (*e.g.*, Factories producing the same kind of commodity), but differ from one another in some aspect which can be measured (*e.g.*, producing different grades of the same article, or different sizes of the same commodity), they together form a Group, and are capable of being tabulated in grades.

(b) Classes.

When the characteristics of the examples are different, but such difference is not capable of being definitely measured, and thus need a separate description (*e.g.*, different methods of production), special classification becomes necessary, and these become Statistical Classes.

(c) Series.

When records are regularly made at intervals of time of the numbers of examples which fall into either or both of the above divisions, the result is a Statistical "Series."

A series of Statistical Data may fall into one of two classes, viz.:

1. *Continuous Series.* This happens when the data is known to fall within certain limits, such as happens with natural objects of all kinds, but the items of which can never be *exactly* measured, *e.g.*, when we are informed that there are 200 examples which exceed size 10 but do not exceed size 15.

2. *Discrete Series.* In this case the items of data are capable of exact measurements such as happens when dealing with wages or profits. In recording and tabulating such date gaps in the magnitudes may be expected, and the continuity of the table is thus broken, *e.g.*, there are 150 men employed who each earn £3 10s. per week.

§ 12.—Summary.

Statistical data are incapable of proving anything, as they are entirely inanimate. It is when the figures are used in an unscientific manner that they can be made to bolster up fallacious statements. All statistics are open to criticism until the psychology behind them has been analysed. Nothing is a Statistical fact which—

(*a*) Does not give the complete information relating thereto;

(*b*) Is not truly representative of the whole data;

(*c*) Is subject to personal bias, however innocent such bias is;

(*d*) Is embellished by personal opinion;

(*e*) Has not been carefully verified and checked.

In checking the conclusions drawn from a Statistical statement—

(*a*) Ascertain that all the data is pertinent to the investigation in hand.

(*b*) That all data conforms to the definition laid down.

(*c*) That all data is homogeneous.

(*d*) That all data being used for the purposes of comparison is strictly comparable.

(*e*) The matter should be approached with an open mind, and not with a previously formulated idea of what will be found.

(*f*) If the result revealed does not agree with what might be expected from a study of the averages or tendencies it should not be rejected without very careful investigation, as it may be an abnormal item which needs inclusion. (This is fully dealt with in the chapter on averages and types.)

(*g*) Where monetary values are used they should be checked by a study of the quantities and qualities of goods involved.

(*h*) All possible and probable causes for phenomena shown should be investigated, with a view to ascertaining which is or are responsible, and the exact effect of each where several are present.

SYNOPSIS TO CHAPTER III.

ACCURACY AND APPROXIMATION.

§ 1.—Accuracy.
- (a) Standard of Accuracy.
- (b) How degrees of Accuracy may be shown.

2.—Approximation.
- (a) Methods of Approximation.
  - i. Approximation to nearest number or specified decimal place.
  - ii. Approximation by discarding figures.
  - iii. Approximation by adding figures.
- (b) Results of Approximation upon Percentages.

3.—Errors.
- (a) Measurement of Errors
  - i. Absolutely.
  - ii. Relatively.
- (b) Classes of Errors.
  - i. Unbiassed Errors.
  - ii. Biassed Errors.
- (c) Estimating the Error.
  - i. When the error is Unbiassed.
  - ii. When the error is Biassed.
  - iii. In an Estimated Average.

4.—Fallacies in the Use of Percentages

5.—Ratios and their Use.

6.—Effect of Errors on Ratios.

7.—Decimals and their Advantages.

CHAPTER III.

ACCURACY AND APPROXIMATION.

§ 1.—Accuracy.

It has already been stated that Statistics is a science of estimates, in other words an application of the theory of probabilities. We obtain the time taken to complete a certain piece of work from past records and we base future estimates for similar work upon our past experience of how long similar work took. We consider it probable that we shall require a certain amount of material and labour, and make provision accordingly. We estimate the quantity of wheat, or oats or barley produced annually because there is no means of measuring either the amount produced or consumed. The number of people who are present at a function is estimated more or less correctly, but no one would know exactly how many persons were present. The company promoter estimates the amount of profit which he anticipates the Company will be able to make, basing his figures upon the past experience of the business which is being acquired, or else upon the results obtained by other companies conducting similar business. Obviously, however carefully past experience is studied, we cannot expect to estimate future results with absolute accuracy, though we can expect to get within a small margin of the actual

result. In commercial Statistics, where the human element plays so large a part, we may well expect to find a greater discrepancy between our estimates and the result than in social and physical statistics, for many factors over which the business man has no control may affect the demand he has anticipated. In any estimate, therefore, we may expect to find errors, and this expectancy becomes greater when we are using approximated figures, for here our data are admittedly incomplete at the commencement of our investigation.

(a) Standard of Accuracy.

Bearing in mind that absolute accuracy is an impossibility whenever it becomes necessary to use measurement, we must lay down a definite standard of accuracy for each estimate, and endeavour to maintain it throughout similar statistical work. We shall naturally alter our standard to fit the case being reviewed. For instance, if it becomes necessary to ascertain the quantity of coal contained in a large dump of that very essential commodity, the contents of the dump are usually estimated after it has been measured, and this estimate if expressed in tons will probably suffice for all ordinary purposes, since in so large a quantity a few pounds or even hundredweights will make but little difference. But if a chemist is estimating the quantity of poison in his stock, then such estimate would need to be correct to the nearest grain, as otherwise a serious mistake might be made. When the exact object for which the estimates are required, or when the known object entails further computations being made, the degree of accuracy present must be indicated, as otherwise the worker is not aware to what extent he can rely on the data.

(b) How degrees of Accuracy may be shown.

Degrees of Accuracy may be shown in various ways, and the following are some of them :—

(*a*) The quantity of coal is 1,000 tons in round numbers.

(*b*) The quantity of coal is 1,000 tons, plus or minus an amount not exceeding half a ton.

(*c*) The quantity of coal is $1{,}000 \pm \cdot 5$ tons. (This is the mathematical way of stating (*b*), but otherwise is the same.)

(*d*) The quantity of coal is between 999·51 and 1,000·5 tons.

(*e*) The quantity of coal is 1,000 tons correct to ·05 per cent., or ·5 per mille.

In this case, as the difference is small, we might rely upon round numbers as in (*a*).

§ 2.—Approximation.

Since numerous digits are confusing to the eye, and also to the mind, it is usually found advantageous when dealing with numbers composed of many figures to express them in round numbers, even when the actual numbers are themselves available. The smaller number of digits simplifies comparison and renders calculation easier, and can generally be used unless the difference between the actual numbers is so small as to preclude the use of approximations.

(a) Methods of Approximation.

(i) Approximation to nearest number or specified decimal place.

An example will illustrate the effect of such approximation. Columns (*a*), (*b*) and (*d*) of the following

Table, comprising figures relating to the working of the Railways of the United Kingdom for a period of five years and show (*a*) the Total Gross Receipts, (*b*) the Total Working Expenditure, and (*d*) the proportion which the latter bears to the former in percentages calculated to the nearest whole number. Column (*c*) has been added to show the percentage calculated correct to the second decimal place. When such calculations as these are used the last digit in each case is supposed to be correct, *i.e.*, it is the nearest figure. For example, in column (*c*), Table A below, the figures in the first line should read 62·617 +, but a more correct result will be obtained when approximating by using 62·62 than 62·61, since the exact number is nearer to 62·62 than to 62·61. It is often advisable to carry the calculation to two or more places beyond the figure to which correctness is desired, as, for instance, when the percentage reads 6·509. If the percentages were being stated correct to the nearest 1 per cent., and only the figure in the first decimal place has been calculated, the observer will not be quite sure whether to read this as 6 or 7 per cent., nor will the second figure in this case assist him in arriving at a decision, whereas the third figure being 9 will cause the approximation to be made to 7 per cent., for if the third figure be 9, then on approximating to the second figure this becomes 1, and the result will then read 6·51, which being over one-half will be approximated to the next number above. Thus the number 52·2500001 when approximated to the first decimal place will become 52·3, for whenever a fraction exceeds half, no matter by how small a margin, it should be counted as a whole number. All fractions under half should be discarded (*i.e.*, 52·34995 would become 52·3 when written correct to the first decimal place).

When the fraction is exactly one-half it should be retained or dropped at the discretion of the compiler, though when such examples are at all numerous half of the examples should be retained, and the other half discarded, thus ensuring a more correct result in the total.

TABLE A.

TABLE SHOWING TOTAL GROSS RECEIPTS AND TOTAL WORKING EXPENDITURE OF RAILWAYS IN THE UNITED KINGDOM FOR FIVE YEARS.

	(a)	(b).	(c)	(d).
Year	Total Gross Receipts	Total Working Expenditure.	Proportion of Working Expenditure to Gross Receipts, calculated correct to Second Decimal Place.	Proportion of Working Expenditure to Gross Receipts correct to the nearest 1 Per Cent.
	£	£	Per Cent.	Per Cent.
1	139,451,429	87,320,550	62·62	63
2	139,098,365	88,173,232	63·39	63
3	144,860,250	93,378,872	64·46	64
4	154,467,914	102,521,053	66·37	66
5	168,721,469	115,993,894	68 75	69

The figures in columns (*a*) and (*b*) are of such a magnitude that they are cumbersome to handle and difficult of comparison. This latter difficulty has been overcome by the insertion of the percentage columns, but even in order to obtain these percentages fairly long calculations are necessary, and mistakes are possible. Comparison may be rendered easier and calculations simplified were we to take round numbers, *i.e.*, approximate the numbers to, say, the nearest thousand, hundred thousand, or even to the nearest million, or, if preferred, we can state the figures in

millions correct to the third, second or first place as the case may be, without appreciably affecting the proportion, as is shown in the following Table, compiled from Table A, but approximated to the nearest thousand :—

TABLE B.

TABLE SHOWING DATA IN TABLE A ABOVE, APPROXIMATED TO NEAREST THOUSAND.

Year	(*a*).	(*b*).	(*c*).	(*d*).
	£	£	Per Cent.	Per Cent.
1	139,451*	87,321	62 62	63
2	139,098	88,173	63·39	63
3	144,860	93,379	64·46	64
4	154,468	102,521	66·37	66
5	168,721	115,994	68·75	69

ii. APPROXIMATION BY DISCARDING FIGURES.

Another method of approximation sometimes used, is that of discarding certain digits entirely and so arriving at a round number of the units specified. The approximation is thus made to the hundred thousand or million *under* the actual numbers, such as reading 123,126,542 and 123,946,789 as 123 millions in each case ; or, if we are dealing in monetary values, disregarding the shillings and pence and taking £93 2s. 6d. and £93 19s. 11d. each as £93.

iii. APPROXIMATION BY ADDING FIGURES.

In this case the approximation is carried out by taking the number to the next specified unit *above* instead of below the number. Thus 159,463,121 and 159,872,642 would each become 160 millions when

* If no information as to how far the figures were correct were given in the heading of such a Table, this item would be shown as 139,451°°°, thus indicating the degree of correctness of the figures.

approximating in millions. Similarly, £19 1s. 3d. and £19 19s. 2d. would each appear as £20 if we were approximating in pounds sterling.

In all cases where approximated figures are used, there is, of necessity, a difference (often of some magnitude when approximating to units of high denominations, *e.g.*, millions) between the actual number themselves and the estimate (*i.e.*, the approximated number), and care must therefore be exercised in using such numbers. While approximation has but little effect upon percentages, yet if these numbers be used in calculations entailing multiplication and/or division; the raising of powers of the original number; or the extraction of a root, the divergence is greatly affected, and the error present is multiplied or divided as the case may be. Thus approximated figures should be avoided when exact results are required from such calculations.

(b) Results of Approximation on Percentages.

A comparison of the two final columns of the Table B with those of Table A will show that the approximations have not made any difference in the proportions which the figures in column (*b*) bear to those in column (*a*), thus demonstrating that in such large numbers an approximation to the nearest thousand will not in any way affect any conclusions which may be drawn from the smaller numbers of figures now used. Table C shows the same data approximated to the nearest million, and a study of the percentage calculations will show that though the percentages vary in all cases in the decimal places, yet when they are taken correct to the nearest 1 per cent., only the fourth year shows any divergence from the results obtained when the actual

figures were used. It should, however, be noted that percentages correct to the nearest 1 per cent. are usually considered sufficient for comparative purposes, since working to several places of decimals not only involves much work, but does nothing to simplify comparison, which is the main advantage of using percentage calculations.

TABLE C.

TABLE SHOWING DATA IN TABLE A ABOVE, APPROXIMATED TO THE NEAREST MILLION.

Year.	(*a*).	(*b*).	(*c*).	(*d*).
	£	£	Per Cent.	Per Cent.
1	139	87	62·59	63
2	139	88	63·31	63
3	145	93	64·14	64
4	154	103	66·88	67
5	169	116	68·64	69

§ 3.—Errors.

The divergence between the actual number and the estimate which is made either by approximation or by any other method, is called by Statisticians an **Error,** and care must be taken not to confuse this use of the word with that to which it more usually applies, *i.e.*, a mistake.

(a) Measurement of Errors.

Errors are measured in one of two ways, *viz.* :—

i. ABSOLUTELY.

If we estimate the population of a town to be 99,900, whereas actually the population is 100,000, then the difference (in this case 100) between the estimate and the exact number is the Error in our estimate. Such an Error is known as an "**Absolute Error.**"

ii. Relatively.

If we express the absolute error as a ratio of the estimate, we obtain what is called a "**Relative Error.**" Thus, though the absolute error in the case just quoted is 100 in an estimate of 99,900, the Relative Error is $\frac{100}{99,900}$ or ·001001. The Relative Error is sometimes expressed as a percentage error, and were this done in the above case the error would be said to be ·1001 per cent. of the estimate.

It will easily be appreciated that the Relative Error is a much more reliable guide to the position than is the absolute error. For example, if it be estimated that there are 100 people at a function, but there are actually only 99 then there is an absolute error of 1, and a similar absolute error would be disclosed where the estimate is 100,000 and the actual number is 99,999. While the absolute error is the same in both cases, the relative error is $\frac{1}{100}$ in the first case and $\frac{1}{100,000}$ in the second, *i.e.*, the effect of the second error is only one one-thousandth of that in the first.

(b) Classes of Errors.

Errors fall into two classes, *viz.* :—

i. Unbiassed Errors.

In the case where the estimate may be either above or below the actual number, such as happens where approximation is carried out to the *nearest* complete unit decided upon (*i.e.*, a hundred, a thousand, etc.), the estimated number may be either greater or smaller than the actual number, and when this happens the error is said to be an "**Unbiassed**" or "**Compensating**" error, because the error may be in either direction, and thus an error in one number in one

direction tends to be offset by an error in another number in the opposite direction, but there must, of necessity, be a large number of items under review before we can expect the Relative Unbiassed Error to become a negligible quantity.

II. BIASSED ERRORS.

If our estimates are all known to be either over or under the actual figures, as happens when we approximate to the next complete unit above or below the exact number, the Error is a "**Biassed**" or "**Cumulative**" error, *i.e.*, the errors are all on one side, and consequently in totalling a number of items the error becomes enlarged. The greater the number of items the greater the total error will be, as will be seen from the following Table, the numbers therein having been chosen at random, and on no preconceived plan:—

TABLE D.

TABLE SHOWING RESULTS OF BIASSED AND UNBIASSED ERRORS.

(*a*).	(*b*).	(*c*).	(*d*).	(*e*).
Actual Numbers.	Numbers correct to the Nearest 1,000.	Unbiassed Absolute Error.	Numbers correct to next 1,000 under	Biassed Absolute Error.
50,750	51	+ 250	50	— 750
76,305	76	— 305	76	— 305
82,805	83	+ 195	82	— 805
40,602	41	+ 398	40	— 602
62,643	63	+ 357	62	— 643
102,486	102	— 486	102	— 486
37,284	37	— 284	37	— 284
25,521	26	+ 479	25	— 521
106,972	107	+ 28	106	— 972
585,368	586	+ 632	580	— 5368
Average of Items: 65,040·9	65,111·1	..	64,444·4	
Relative Errors:	..	·001078	..	·009255

It will be seen from this Table that in the case of Unbiassed Errors the Absolute Error in the Total of a number of items, each of which is subject to the same unbiassed absolute error, will tend to increase with the number of items, but the relative error present in the total will diminish with the number of items. In the above Table the Absolute Error in the first item, as shown in column (*c*), is 250, and the estimate is 51^{000} [column (*b*)], and the Relative Error is therefore $\frac{250}{51,000}$, or ·004901. In the totals the Absolute Error is 632 in an estimate of 586^{000}, which gives us a Relative Error of $\frac{632}{586,000}$, or ·001078, a greatly decreased Relative Error. It will also be seen that the Unbiassed Absolute Error is almost negligible when we consider it in the light of the magnitude of the figures of the estimate, whereas the Biassed Absolute Error is very material.

(c) Estimating the Error.

i. When the Error is Unbiassed.

If the exact figures composing the number were unknown, and the only information at our disposal was that the total of the items when approximated to the nearest thousand was 586,000, it would be very desirable to know what possibility of error existed in the estimate. Each of the items comprised in the total may have an Absolute Error, and this error may be anything from 0 to 499, since the number may be an exact number of thousands, and the approximate and exact numbers thus coincide, or any number over the thousand and up to the next five hundred would cause an error varying with the number by which it exceeds the thousand, similarly with the numbers over the five hundred and under the next thousand. The most

likely error, however, will be the average of these numbers 0 to 499, since it is the average type of anything which we meet with most frequently, while the extremes are seldom in evidence. The average of the successive numbers 0 to 499 will be found to be approximately 249·5, or in round numbers 250, and this latter figure is sufficiently correct for ordinary purposes. To obtain the best estimate of the **Absolute Error** in a Total of a number of items which have an Unbiassed Error we *multiply the average Absolute Error of the items by the square root of the number of items.* If we apply this to column (*b*) of the above Table D, we shall get $250\sqrt{9} = 250 \times 3 = 750$, so that were the actual numbers and the total unknown we could write our estimate as $586{,}000 \pm 750$, since the error may be in either direction. As we know the exact numbers in this case we are able to compare the estimate with the exact quantity. The actual error as shown in column (*c*) is 632, and our estimate of 750 is therefore fairly good. It will be found in practice that the estimated error and the actual one do not often agree, and exact results cannot therefore be expected. However carefully and scientifically we work we cannot lose sight of the fact that the estimated error is based on the probability that the average Absolute Error of the items is the most likely error to be made.

If we have an estimate of the total of a number of items each of which is subject to the same Unbiassed Error, and it is desired to ascertain what is the possible **Relative Error** of such estimate, we can obtain the best possible estimate of the Relative Error if we *multiply the average Absolute Error of the items by the square root of their number, and divide by the estimated total.* From the above Table we know the estimated total to be 586,000 and the average Absolute Error to be approximately 250, while the number of items is 9.

The estimate of the Relative Error in such estimated total would therefore be :—

$$\frac{250\sqrt{9}}{586,000} = \cdot 00128$$

From the actual results which are known to us we can compare this result with that actually correct. This latter is found to be ·001078, so that our estimate is not far from correct. We could therefore write our estimate as 586,000 correct to ·128 per cent.

ii. When the Error is Biassed.

When we are approximating to the next complete unit below, as in column (*d*) of the above Table D, the total of the errors found in the items will be the Absolute Error of the total, and this will be observed by studying column (*e*). In this particular case the error in each item may be any number between 1 and 999*, since all figures between these two extremes are equally likely to have been omitted. As the error is a biassed one due to the fact that we are discarding digits, it will increase with the number of items under observation. It is obvious that the minimum error possible is 1 × *the number of items in the Table*, and the maximum error is 999 × *the number of items*. The most likely error, however, will be the average error possible, and hence the best estimate we can make of the **Absolute Error** in the total will be *the average Absolute Error of the items multiplied by the number of items.* If the average of the consecutive numbers 0 to 999 be calculated, it will be found to be approximately 499·5 (or, say, 500 in round numbers). Applying this fact and the above rule to the Table D above, we find that the total appearing in column

*The error in this case has been taken in terms of whole numbers, but if fractions be considered the error in each item could be anything between, say, ·001 and 999·999.

(*d*) to be in error by an estimate of $499{\cdot}5 \times 9 = 4495{\cdot}5$ (or if round numbers be taken $500 \times 9 = 4{,}500$), whereas a comparison with the actual figures which are available shows the actual error of the total to be 5,368. Though the estimated error in each case is not likely to be correct, this method of computing the possible error in an estimate is very useful when it is impossible to obtain the exact figures, and we must be content to work on estimates. In the absence of any exact data, therefore, we could say that the total of the nine items when approximated to the next thousand under would be 580,000 + 4,500, as the estimate is "short" or below the exact number in each case, while if the approximation had been made to the next thousand above, the estimate would need to be written 580,000 - 4,500, as in this case the estimate is definitely known to be in excess of the actual number.

If we desire to obtain an estimate of the **Relative Error of the total** when the error is a biassed one, we *multiply the average Absolute Error of the item by the number of items and divide by the estimate.* Applying this to the above case (Table D), the average Absolute Error of the items in column (*d*) is 500 in round numbers, the number of items is 9 and the estimate is 580,000 then the estimated relative error is

$$\frac{500 \times 9}{580{,}000} = {\cdot}0077586$$

whereas the correct relative error is ·009255.

iii. Errors in an Estimated Average.

Frequently we find that instead of the actual numbers and totals we have only an estimate of the average of a certain number of items which are known to be correct to a defined figure. Thus the average of the items in column (*b*) in the Table above is 65,111·1 when calculated correct to the first decimal place, and we desire

to ascertain the approximate variation which there is likely to be from this average. To obtain the most probable variation we *divide the average Absolute Error of the items* (*i.e.*, in this case 250) *by the square root of the number of items* (9), and we thus get: $\frac{250}{\sqrt{9}} = 83 \cdot \dot{3}$, so that we can say the average of the items is estimated to be 65,111 ± 83. Again, as we have the actual results available we are able to compare the estimate with the actual figures.

The average of column (*a*) is 65,040·9 correct to the first decimal place, and hence the error in the averages of columns (*b*) and (*a*) is actually 65,111·1 − 65,040·9, or only 70·2. It will at once be seen from this result that the Absolute Error of the total when divided by the number of items will give us the Absolute Error of the average of the numbers comprising the total (*i.e.*, 632 ÷ 9 = 70·2), and also **that the Unbiassed Errors tend to disappear in the averages.**

The **Relative Error of an average,** when the error is an Unbiassed One, can be found by dividing *the average Absolute Error of the items by the average of the numbers as estimated, and also by the square root of the number of items*; *e.g.*, taking the figures in Table D above, the Relative Error of the average of column (*b*) may be estimated as:—$\frac{250}{65{,}111 \cdot 1 \sqrt{9}} = \cdot 00128$, a result which agrees with the estimated Relative Error of the total, and it will be apparent, therefore, that the Relative Error of the total and the Relative Error of the average should agree.

§ 4.—Fallacies in the Use of Percentages.

Whenever percentages are used it is essential that the basis of the calculation of the percentage be given,

so that there can be no misunderstanding. In the following statement this has not been done, and consequently two results can be obtained from the data, but these results vary greatly:—

> "Wages were raised 10 per cent., lowered 15 per cent., raised 25 per cent., and lowered 20 per cent., and then raised 15 per cent. in certain years. What was the change in wages over the whole period?"

What was the basis on which the percentages were calculated? Was it on the original wages, or on the rate of wages ruling at the date the change was made?

Case I. If the percentages were based on the original wages, then the change over the whole period would be obtained as follows:—

Let the original wages = 100.

Then

(*a*) the 1st change is an increase of 10 per cent., then

$$\text{wages} = \frac{110}{100} \text{ of } 100 = 110.$$

(*b*) The 2nd change is a decrease of 15 per cent., then

$$\text{wages} = 110 - \frac{15}{100} \text{ of } 100 = 95.$$

(*c*) The 3rd change is an increase of 25 per cent., then

$$\text{wages} = 95 + \frac{25}{100} \text{ of } 100 = 120.$$

(*d*) The 4th change is a decrease of 20 per cent., then

$$\text{wages} = 120 - \frac{20}{100} \text{ of } 100 = 100.$$

(*e*) The 5th change is an increase of 15 per cent., then

$$\text{wages} = 100 + \frac{15}{100} \text{ of } 100 = 115.$$

So that the change of wages over the whole period has been an increase from 100 to 115 per cent., or a change of 15 per cent. on the basic rate.

Case II. If the changes had been based on the rates of wages ruling at the time of the change, the following would be the result :—

Again let the original rate of wages = 100.

Then

(a) The 1st change is an increase of 10 per cent., then

$$\text{wages} = \frac{110}{100} \text{ of } 100.$$

(b) The 2nd change is a decrease of 15 per cent., then

$$\text{wages} = \frac{85}{100} \text{ of } (a).$$

(c) The 3rd change is an increase of 25 per cent., then

$$\text{wages} = \frac{125}{100} \text{ of } (b).$$

(d) The 4th change is a decrease of 20 per cent., then

$$\text{wages} = \frac{80}{100} \text{ of } (c).$$

(e) The 5th change is an increase of 15 per cent., then

$$\text{wages} = \frac{115}{100} \text{ of } (d).$$

Then final wages =

$$\frac{115}{100} \text{ of } \frac{80}{100} \text{ of } \frac{125}{100} \text{ of } \frac{85}{100} \text{ of } \frac{110}{100} \text{ of } 100 = 107{\cdot}525.$$

So that wages have risen in the period under review from 100 to 107·525, a difference from the previous result of 7·475. Such variations in results would not arise were proper information always supplied.

Percentages, while useful to enable comparison to be made, should only be used when the factors which are expressed in percentage form are themselves

comparable, as if such is not the case any derivative is also not comparable. In dealing with such matters as profits, for example, it is frequently contended that by using percentages we get a more correct result than that which would be obtained by giving the actual figures for profit. This may be true in many cases but is not correct when any change has taken place in the magnitude of the Capital involved, as has been abundantly shown in the recent results of many Companies which increased their capital. Let us assume for the sake of example that a Company with an issued and paid up capital of £250,000 has consistently paid dividends of 8 per cent. per annum. Ignoring any questions of appropriations to Reserve, or the balance of profit usually carried forward by Companies, their profits, therefore, amounted to no less than £20,000 per annum. Supposing for various reasons the Company finds it expedient to increase its Capital by 100,000 fully paid shares of £1 each: while it is true that such additional capital may earn the same rate of profit as that originally invested, yet it will probably take some time before it becomes fully remunerative, as development is usually steady rather than sudden. If in the first year in which the additional capital is in the business the profits amount to £24,500, an increase of £4,500 over the results of the previous year, progress has certainly been made, but the percentage of profit to capital is only

$$\frac{24{,}500}{350{,}000} \times 100 = 7\%.$$

Thus, if only a percentage table of profits for five years were shown as follows:—

First Year	8%
Second ,,	8%
Third ,,	8%
Fourth ,,	8%
Fifth ,,	7%

we should assume that profits had declined instead of having shown an increase, which would have been revealed were the actual profits shown. In such cases as this it is essential to show the whole of the data if fallacies are to be avoided. This could be done as follows:—

Year.	Capital.	Profits.	Rate of Dividend.
	£	£	
1	250,000	20,000	8%
2	250,000	20,000	8%
3	250,000	20,000	8%
4	250,000	20,000	8%
5	350,000	24,500	7%

§ 5.—Ratios and Their Use.

Some Statisticians strongly recommend that ratios be used instead of percentages, and this would certainly do much to eliminate mistakes which arise when percentages are used and the differing bases are not specified, since we should then state that Wages had increased in the ratio of 100 to 115 (100 : 115) rather than that Wages had increased by 15 per cent. It is advisable whenever possible to dispense with decimal points, as these are frequently either omitted or misplaced, so that in the case (2) above we should find it necessary to state the ratio as 100,000 : 107,525, rather than as 100 : 107·525.

It is not wise to trust always to percentages for comparative purposes, as the following little anecdote will amply demonstrate. Two army medical officers were discussing the efficiency of inoculation against typhoid, and one of them stated that of those patients he had inoculated 50 per cent. had subsequently contracted the disease. The other was much astonished at this statement, for, said he, "My experience

is the reverse of yours, for of those cases which I inoculated only one-tenth of 1 per cent. subsequently were victims to the fever." The two opinions were so widely divergent that further inquiries were made, and it transpired that whereas the first medico had only inoculated two men, of whom one had subsequently sickened with the fever, the second officer had treated no fewer than 1,000, and only one of these had contracted the disease. Apparently, on the face of the statements, they were capable of comparison, since they were expressed as percentages, yet in reality no comparison was possible owing to the varying base of calculation. If ratios had been used, no misunderstanding would have been possible, since they would have been 1 : 2, and 1 : 1,000, these figures being obtained by reducing the actual numbers to the simplest form.

Other examples of a similar nature will readily occur to the reader, but one further illustration may perhaps be given with advantage. Results of various kinds are often given in percentage form, particularly in advertisements. A coaching institution may state that its percentage of passes in a certain examination is 100 per cent., while another has a percentage of only 95. The first institution would appear to be better for a prospective candidate than the second, but supposing the first named had only prepared two candidates, both of whom had passed, while the second had coached two hundred, of whom 190 had been successful, the opinion would of necessity be reversed. Here again ratios would have prevented fallacious conclusions, and it may therefore be taken as an axiom that percentages alone are insufficient ground for comparison, unless the data upon which they are based is known to be both homogeneous and comparable.

Care must be taken even in the use of ratios or inaccurate conclusions may be drawn, particularly when they invite comparison between unrelated figures. Thus, if in an examination there were 500 candidates of whom 300 are successful the ratio of passes to failures is 300 : 200 or 3 : 2. Supposing a coaching college had prepared 300 of these students of whom 200 passed, the ratio of successes to failures is 2 : 1, whereas of the other 200 candidates 100 must have been successful which gives a ratio of 1 : 1 ; thus making it appear that the college in question obtains twice as good results as all others combined. This is obviously fallacious, as some candidates may not have been coached at all, and hence do not stand the same chances of success as those properly prepared. The only proper method is to show the ratios of all the coaching institutions if we desire to ascertain which has the best results.

§ 6.—Effect of Errors on Ratios.

If for the purposes of comparison we are preparing ratios between two distinct groups of data, it is essential that we should know what effect any errors in the estimates we use will have upon such ratios. In the case of Unbiassed Errors it will be found that they do not usually affect the ratios at all, while Biassed Errors tend to disappear in the ratios. This may be illustrated in the following figures :—

	Actual Numbers to be Compared.		Numbers Approximated to Nearest 1,000.		Numbers Approximated to Next 1,000 below.	
	564,584	1,693,752	565	1,694	564	1,693
Ratios:	1 : 3		1 : 2·99999		1 : 3·002	

§ 7.—Decimals and Their Advantages.

When calculations of any kind become necessary, and fractions form a part of the numerical data with which the observer is dealing, such fractions should always be expressed as decimals, for as such they are not only easier to read, but also simplify calculation in so far that Logarithms, or the Slide Rule, can be used if the work is long or intricate. Useful as Logarithms are, it must not be overlooked that in all cases where they are used the answer is only correct to the last figure of the answer, or to the last digit, when succeeded by one or more ciphers, and thus the figures cannot be relied upon when further computation is necessary.

When approximations involving the use of decimals are used it is essential that, if the figure to which correctness has been calculated is a cipher, such cipher should be shown in the same manner as if the figures were a digit. For instance, if the heights of buildings are being estimated, and one is expressed as being approximately 100 feet in height, the actual height might be anything between 99·51 feet and 100·49 feet (or 99·5 and 100·5), whereas if the height be shown as approximately 100·0 feet, then the measure of variation which is possible is between 99·95 feet and 100·05 feet, thus allowing a narrower margin of error in the estimate.

SYNOPSIS TO CHAPTER IV.

TYPES AND AVERAGES.

§ 1.—USE OF AVERAGES.

2.—THE ARITHMETIC AVERAGE.

- (a) The Simple Arithmetic Average.
  - i. Short Method of Computing.
  - ii. Advantages of the Simple Arithmetic Average.
  - iii. Disadvantages of the Simple Arithmetic Average.
  - iv. Commercial Uses of the Simple Arithmetic Average.
  - v. Moving Average Standard
    - (1) Uses of Varying Average Standards
    - (2) Short Term Fluctuations.
    - (3) Uses of the Trend.
  - vi. Progressive Average.
- (b) The Weighted Arithmetic Average.
  - i. Definition of the Weighted Arithmetic Average.
  - ii. The Weights to be used.
  - iii. When Weighted Averages must be used.
    - (1) When classes of the same group contain a varying number of items.
    - (2) When proportions change.
  - iv. Commercial Uses of the Weighted Arithmetic Average.
- (c) The Descriptive Average.

3.—THE GEOMETRIC AVERAGE.

4.—THE MEDIAN.

- (a) Arraying the Data
- (b) Frequency Tables.
- (c) Rule for calculating the Magnitude of the Median
- (d) Quartiles.
- (e) Deciles and Percentiles
- (f) Advantages of the Median.
- (g) Disadvantages of the Median.
- (h) Uses of the Median.

5.—THE MODE.

- (a) Method of Locating the Mode.
- (b) Advantages of the Mode
- (c) Disadvantages of the Mode.
- (d) Commercial Uses of the Mode.

CHAPTER IV.

TYPES AND AVERAGES.

Whenever we desire to form a standard unit for the purposes of comparison, such standard should always be obtained by eliminating as far as possible, or at least nullifying the effect of, any abnormal fluctuations which may have taken place, and thus be present in the data which it is desired to review. This may best be done by the judicious use of an Average which is obtained by calculation, or a statistical Type which is a selected example, since, by combining those years which show extraordinary phenomena in either direction with those of normal years and obtaining an average spread over a long period, we eliminate to a very large degree the effects of an abnormal period, and obtain a more useful standard for comparison, since it must be recognised that abnormalities are bound to occur from time to time. Moreover, when a long term is used abnormalities in one direction tend to be offset by those in the opposite direction, and consequently a more correct standard is obtained.

§ 1.—Use of Averages.

(*a*) Averages can be used to give us a definite idea of some large group or groups of numerical data, and so enable us to compare different groups and classes

which are of such magnitude in themselves that comparison unit by unit would either be impossible, or take too much time. (*b*) By the careful use of averages, we can often obtain a reliable guide to a complete group or class by the use of samples. (*c*) Averages become compulsory when we desire to obtain an idea of the mathematical relationship existing between different classes or groups.

§ 2.—The Arithmetic Average.

In Statistics we use various kinds of averages, each of which is particularly applicable to certain fields of investigation, but some of these averages are really types. The most common kind of average, and one which has so far been referred to in this manual, is the **Arithmetic Average.** There are three main variations of this Average, viz.: (*a*) The Simple; (*b*) The Weighted; and (*c*) The Descriptive.

(a) The Simple Arithmetic Average.

The sum or total of a number of items is known as the AGGREGATE. *The Arithmetic Average is obtained by dividing the aggregate of a series of items by the number of items in the series, i.e.*, the average of the numbers 3, 4 and 5 is $\frac{3+4+5}{3} = 4$, a computation which involves no particular mathematical skill.

i. SHORT METHOD OF COMPUTING THE SIMPLE ARITHMETIC AVERAGE.

When, however, we are dealing with a large number of items, all of which are of nearly the same magnitude, the following "Short-Cut" method of finding the

Simple Arithmetic Average may frequently be used with advantage :—

TABLE E.

EXAMPLE OF SHORT METHOD OF COMPUTING ARITHMETIC AVERAGE.

Items to be Averaged.	Assumed Average.	Deviations of the Actual Numbers from the Assumed Average
1684	1680	+ 4
1680	1680	0
1686	1680	+ 6
1675	1680	— 5
1673	1680	— 7
1670	1680	—10
1677	1680	— 3
1676	1680	— 4
1682	1680	+ 2
1683	1680	+ 3
Sum of the deviations ..	..	— 14

$$\text{Average deviation} = \frac{-14}{10} = -1\cdot4.$$

Then average of the items = 1680 + (— 1·4) = 1678·5

Explanation: Take an arbitrary figure and assume this to be the average of the numbers under review, and then tabulate the deviations of the actual numbers from such assumed average. These deviations will be added, paying due regard to the signs which precede them, and will then be divided by the number of items. The result will be added to the assumed average, and the answer will give us the true average. Obviously the arbitrary figure should not be too far from the actual average, but this is easily obtained by inspection as in the example.

ii. ADVANTAGES OF THE SIMPLE ARITHMETIC AVERAGE.

(1) It is very easy of computation.

(2) It is understood by everyone, and the method of arriving at it is common knowledge, a fact

which renders it very useful when presenting averages to people possessing no special mathematical knowledge.

(3) There is no need to arrange the data in any special order, as is necessary with some Statistical types.

(4) It may be obtained when only the number of the items and their aggregate are known, and no details of the various items are available. If, for instance, we know the amount raised by taxation in any country, and the population of that country, we can calculate the average amount of Tax contributed per head of the population, although we do not know the actual amount paid by any one individual.

(5) If we know the number of items as well as the average we can calculate the aggregate involved. For instance, given 40 workmen and an average wage per employee of £3 per week we can obtain the weekly pay roll of the works, *i.e.*, £3 $\times$ 40 = £120.

(6) Every example is given due weight in obtaining the average.

(7) If the number of examples used in calculating the average be large, the abnormalities in one direction tend to be offset by those in the other direction, and thus the standard for comparison is more accurate.

(8) It can be used for further calculation when this is necessary.

iii. Disadvantages of the Simple Arithmetic Average.

(1) Unless the data be very simple, the average cannot be obtained by inspection, as can some statistical types.

(2) Any extremes in one direction which are present in the data will affect the average, and though this is desirable in some cases, it is not always so, particularly if the average is to represent the true state of affairs. For example, when calculating the average income of a district where the bulk of the population belong to the artisan class, but in which a millionaire resides, unless the number of artisans was very large the income of the millionaire would greatly affect the average income of the district, and thus give a fallacious average, unless steps are taken to give due weight to the numbers involved.

(3) Its accuracy is impaired unless the whole of the known examples are taken into account.

(4) The result we may obtain may not actually agree with any of the items under review, and may therefore not be truly representative. The following example will clearly demonstrate this. Four workmen earn respectively £4, £5, £6 and £7 per week. The average wage per workman is:

$$\frac{£4 + £5 + £6 + £7}{4} = \frac{22}{4} = £5{\cdot}5$$

but not one of the workman actually receives such a sum.

(5) That when not supported by the actual figures used in obtaining such average it may lead to fallacious conclusions. This leads to a form of abuse of averages, and while the average is correct, yet it is sometimes used to conceal the real position. This can be best illustrated by

taking the case of a Company which in five successive years shows the following profits :—

1st year	..	..	£10,000
2nd year	..	..	8,000
3rd year	..	..	12,000
4th year	..	..	15,000
5th year	..	..	17,500
Total	..	..	£62,500

The average profit for these five years is £12,500 $\left(i.e., \frac{£62{,}500}{5}\right)$ and as the profits for the period show material increases, except in the second year, the obvious inference is that the business is a progressive one, and consequently the Company promoter desirous of launching the Company on the public would show the above figures, as well as the average, the latter being used to indicate what could be expected under average conditions. If, however, the total profits of the five years were the same, but the sequence of profits was :

1st year	..	..	£17,500
2nd year	..	..	£15,000
3rd year	..	..	£12,000
4th year	..	..	£10,000
5th year	..	..	£8,000

then the average profits are still £12.500, but if the actual figures are studied they show that the business is on the downgrade. A company promoter could, however, hide the real facts by giving only the average profits, and thus obtain subscriptions which would not be forthcoming were the whole of the facts known. It was undoubtedly to remedy this that the

Companies Act, 1929, made it compulsory for a certificate of the actual profits for the last three years to be shown in the Prospectus.

iv. COMMERCIAL USES OF THE SIMPLE ARITHMETIC AVERAGE.

The Simple Arithmetic Average lends itself admirably to many types of Commercial Statistics, but care must nevertheless be taken to employ it properly. It has already been pointed out that in comparing any figures relating to various periods care must be taken to see that nothing abnormal occurs in either of the two under comparison at the moment. The Turnover of a business may be affected by a variety of causes. A "Boom" in trade will cause it to increase; a "Slump" will result in a decrease. Inflation of money, or, to put it the other way, a decrease in the purchasing power of money, will, *all other things being equal*, result in a leap upward in the Turnover, while an increase in the purchasing power of money, when other conditions remain unchanged, will, of course, bring about a fall. Increased competition and any change in the Revenue Duties payable upon a commodity (if any) will also affect it, and any one or more of these variations may happen in a very short period. To fix an arbitrary standard by which to measure anything does not allow for any changes of a permanent nature, or even of a semi-permanent nature, which may arise, such as an increase in the Working Capital employed in the business, nor will it do to compare one year with another, since either or both may be subject to abnormalities which are incapable of being separated or assessed. If, however, we make a standard by averaging the results obtained from a number of years we shall have taken steps to eliminate to a very large extent many of the phenomena which would appear to be of a periodic nature, such as a

cycle of trade. In choosing the length of time over which the average is to run care must be taken to see that such cycles of trade are usually covered, and it is therefore advisable to study the figures at one's disposal, in order to see that the period selected will in the normal way cover such cycle. If it is found that cycles of trade take five years, then a five years' average can be used ; but if the cycles take ten years, then it is necessary to take a ten years' average. In ordinary trading concerns a ten years' average will be found to be the most useful. By taking the average over a large number of years, bad years receive an opportunity of being offset by the good ones, or should no good ones be present, to nullify the effects of the bad ones, seeing that the effect of the poor periods is spread over a larger number of years. This would in effect be an Index Number, a subject which will be dealt with more fully later in this manual, and all future years would be compared with this Standard. Though this system has many advantages, it is not perhaps the most suitable for business use, for in business, conditions change from year to year. The working capital of the concern may have been increased, either through an additional issue of Shares or Debentures, or by leaving profits which have been realised in the business and so increasing the interest of the proprietors in the business. All other things being equal, an increase in the working capital at the disposal of a company, or firm, results in better opportunities for purchasing goods, and should logically increase the Turnover. Fluctuations are not therefore confined to outside disturbances, but may be caused by changing conditions within the business, and such changes should be provided for if our standard is to be a useful one.

v. Moving Average Standard.

It would therefore probably be better to make the standard a Moving Average one. This would be done by maintaining the same number of periods, but dropping off the earliest and taking in its place the last trading period, this procedure being repeated year after year. This would enable the Trading of each year to be compared with the ten years' average (or such other average as may be adopted), ending with the previous trading period. The Standard thus obtained would have taken into account all fluctuations due to whatever cause right up to the beginning of the period under review.

(1) *Uses of Varying Average Standards.*

The effect of such Moving Average Standard may perhaps best be illustrated by an example. The following are the figures of the Turnover of a Company for twenty years. The Capital of the Company at the commencement of the period was £10,000, but Reserves amounting to £5,000 had been accumulated by the fifteenth year, and increased to £8,000 by the end of twenty years.

Year	£	Years	£
1	35,000	11	55,000
2	42,000	12	54,000
3	43,500	13	56,500
4	47,000	14	60,000
5	48,500	15	100,000
6	52,000	16	54,000
7	36,500	17	66,500
8	37,500	18	112,500
9	53,000	19	154,000
10	53,500	20	194,000

It will be observed that the Turnover fluctuated from year to year, due to many causes. The reasons of the fluctuations could be ascertained by studying the conditions ruling in the trade itself, and in business generally, during the period under review.

If we worked on the basis of comparing the various years' Turnover with that of the first year, none of the changing conditions would be brought into consideration, for even though we were able to say by how much the Working Capital in, say, the twentieth year exceeded that of the Company in the first, we cannot possibly decide by how much the turnover has increased directly as a consequence of the increased Working Capital. While basing our comparisons on the average of the ten years, 1 to 10, improves the Standard considerably, yet it does not take into account the changing conditions which applied during the second decade, particularly in the growth of the extra Working Capital in the shape of the Reserve. The moving average overcomes this objection, and in addition gives an indication whether the "Trend" of the business is upward or downward. Table F (see p. 79) shows the ratios which the Turnover of the years 11 to 20 bear to the bases obtained by each of the methods outlined above, the ratios being calculated correct to the nearest whole number.

A comparison of columns (*a*), (*b*) and (*c*) is interesting, for while (*a*) shows a fall in the sixteenth year as compared with the preceding years, the fall does not appear so considerable, while the jump upward in the following year is exaggerated owing to the low base used for the comparison. The same phenomena, but to a lesser extent, are also observable in column (*b*), but in column (*c*) the importance of the fall is emphasised as being *below* the standard of the Trend of the business. It might be argued that year 15 was a better year, due to increased Turnover following the known increase in the Working Capital, and that this fact would tend to exaggerate the fall in the following year; but this argument will not hold when we consider that the

Reserves *had been accumulated, and by the fifteenth year had amounted to* £5,000 ; and further, from column (*d*) we see that though the Turnover leapt upward very sharply in year 15, the effect upon the moving ten-year average was not extremely marked, for the average of

TABLE F. *

TABLE SHOWING THE RATIO TURNOVER BEARS TO STANDARDS CALCULATED ON VARYING BASES.

	(*a*)	(*b*).	(*c*).	(*d*).
Year.	Base: Year 1 =£35,000=100.	Base: Average of Ten Years, (1-10) =£44,850=100	Base: Moving Ten-Year Average shown in Col. (*d*)	Ten-Year Moving Average for Col. (*c*).
11	157	23	123	Years 1–10 £44,850
12	154	120	115	Years 2–11 £46,850
13	161	126	118	Years 3–12 £48,050
14	171	134	122	Years 4–13 £49,350
15	286	223	197	Years 5–14 £50,650
16	154	120	97	Years 6–15 £55,800
17	190	148	119	Years 7–16 £56,000
18	321	251	191	Years 8–17 £59,000
19	440	343	232	Years 9–18 £66,500
20	554	433	250	Years 10–19 £76,600

Basis for Comparison for 21 = Years 11–20 = £90,650.

the years 5–14 was £50,650, while for the years 6–15 (the basis of comparison with the year 16) was only

* The arrangement of this Table should be noted, as by placing the information in column (*c*) before that contained in column (*d*) we are enabled to compare the Ratios more easily than would be the case had the information in column (*d*) intervened between that in (*b*) and (*c*).

£55,800, though the Turnover had actually increased from £60,000 in year 14 to £100,000 in the fifteenth year. A further study of column (*d*) reveals the fact that the heavy fall in year 16 did not cause a very heavy sag in the moving ten-year average, for this average continued to advance throughout the ten years under review, and it is this moving ten-year average which shows the Trend of the business turnover.

(2) *Short Term Fluctuations.*

When the average moves upwards it would show that the Trend of the Turnover was increasing steadily during the period, and that the falls which took place in the Turnover from time to time had no real effect on the progress of the business, so that by taking an average over a term of years we actually eliminate those fluctuations which, though having a marked effect at the moment, do not really affect the progress of the business at all. Such fluctuations are called "**Short Term Fluctuations.**" If data were available in relation to the particular trade or to trade in general, it would be possible to compare our ratios with those so published to see whether the business could show a better or a poorer result than the trade as a whole.

(3) *Uses of the Trend.*

The principle which has been enunciated above may easily be employed in other directions. Instead of showing the percentage which Expenses bear to Turnover, which is at present the usual way of comparing the results of a current year with past periods, we could create for expenses a standard similar to that calculated for Turnover, and from this obtain the ratio which the Standard of Expenses bears to the standard

Turnover and compare such ratios with those of the current year. Variations would thus be shown and traced to their source, waste eliminated and excessive expenditure reduced. If the ratio of Turnover to the Standard advanced, while the ratio of current Expenses to the Standard advanced by a similar or a less amount, it would appear that the management was pursuing a sound policy, and *vice versa;* hence, if the standard for expenses was 50 to 100 for Turnover, and in the current year was 70 for expenses and still 100 for Turnover, inquiries would of necessity be made, for expenses must have risen considerably in the period under consideration, and the causes of the rise should be traced to their source, as it may be found that this is due to some condition which is capable of being remedied, or some inefficiency which should be corrected. To carry out such an inquiry properly it may be necessary to analyse the items composing expenses, and compare them with the corresponding items in the standard of expenditure, following the same procedure in each case. With expenses it is essential that the moving Standard be adopted, for many of the component items will fluctuate considerably from year to year, particularly such items as Wages, Salaries, Rates and Taxes.

vi. PROGRESSIVE AVERAGE.

This is a form of average which is occasionally used, and is very useful during the early years of the life of a business and pending the expiration of such a period as it is desired to include in the calculations of a Moving Average Standard. It is a cumulative average, *i.e.*, at the expiration of each period a new average is obtained from the whole of the data then available, and this average is used as the base of

comparison with the next period. In Table F it will be noticed that we used the Ten-Year Average of the years 1–10 as the base of comparison for the year 11. If the business had been established in 1, then to compare the results of subsequent years the Progressive Average might have been used with advantage. The following table shows its compilation from the data shown on page 77.

Year	Profits	Progressive Average			Increase or Decrease* of Years Profits on Average.		
	£	£	s.	d.	£	s.	d.
1	35,000						
2	42,000	38,500	0	0	3,500	0	0
3	43,500	40,166	13	4	3,333	6	8
4	47,000	41,875	0	0	5,125	0	0
5	48,500	43,200	0	0	5,300	0	0
6	52,000	44,666	13	4	7,333	6	8
7	36,500	43,500	0	0	*7,000	0	0
8	37,500	42,750	0	0	*5,250	0	0
9	53,000	43,888	17	9	9,111	2	3
10	53,500	44,850	0	0	8,650	0	0

The first average is obtained as follows:—

$$\frac{35{,}000 + 42{,}000}{2} = \underline{\underline{\pounds 38{,}500.}}$$

The second average is found by adding the profits of the three years and dividing by 3, *i.e.*,

$$\frac{35{,}000 + 42{,}000 + 43{,}500}{3} = \underline{\underline{\pounds 40{,}166 \text{ 13s. 4d.}}}$$

It will at once be realised that such a Progressive Average includes all data up to date, whatever changes may have taken place, but when such a period has been reached that a representative moving average can be obtained, the Progressive Average should be discontinued in favour of the Moving Average, the reason being, that as time advances the earlier data is not truly representative of present conditions, and

* In actual practice the figures showing decrease would be shown in red. As this is not advisable in a book, decreases are indicated by an asterisk.

although to some extent offset by the more recent examples, results in an average which is biassed by early conditions. This is particularly true in the case of a business which has made very rapid progress, and which has from time to time increased the capital employed. Moreover, the total being spread over an ever increasing number of years tends to reduce the effect of modern fluctuations materially. If such a Progressive Average were used for the whole of the data on page 77, the average for the years 1–20 would be £67,750, and this would be the base of comparison for the twenty-first year; whereas the Ten Year-Moving Average gives us a base of £90,650, a figure which is much more in accord with actual conditions than the former.

The addition of a column showing the increase or decrease on the Progressive Average is very useful in showing whether the business is progressing, for if the actual results are in excess of the average, then the Trend will be upward, but when a fall takes place in any period and is accompanied by a fall in the average, and the latter is higher than the actual profits made, then a detailed investigation becomes necessary, in order that the reasons for the position may be discovered, and adequate steps taken to prevent a further retrograde movement.

(b) The Weighted Arithmetic Average.

Though the simple Arithmetic Average can be used in many kinds of Statistical work, it has its limitations, and these are best illustrated by an example.

The employees of a firm number 10,000, and are divided into three groups, as follows :—

(1) 4,000 Skilled workers with an average weekly wage of £5.

(2) 3,500 Semi-skilled workers with an average weekly wage of £4.

(3) 2,500 Unskilled workers with an average weekly wage of £3.

It is desired to ascertain the average wage received per worker. This might, at first glance, appear to be :— $\frac{£5 + £4 + £3}{3}$ or £4, and if this were so, the weekly wage bill of the firm would be £4 × 10,000 = £40,000 ; but, as a matter of fact, the actual pay roll amounts to :—

$$(£5 \times 4{,}000) + (£4 \times 3{,}500) + (£3 \times 2{,}500)$$
$$= £20{,}000 + £14{,}000 + £7{,}500 = \mathbf{£41{,}500}$$

so that our average is clearly incorrect, seeing that one of the advantages of the Arithmetic Average is the fact that the aggregate can always be obtained when we know the average and the number of examples. But if we multiply the average wages of the men in each group by the number of the men in each group, and add the results, and then divide by the total number of workers, we get :—

$$\frac{(£5 \times 4{,}000) + (£4 \times 3{,}500) + (£3 \times 2{,}500)}{10{,}000}$$
$$= \frac{£20{,}000 + £14{,}000 + £7{,}500}{10{,}000} = \frac{£41{,}500}{10{,}000} = \mathbf{£4{\cdot}15}$$
$$= \mathbf{£4\ 3s.}$$

which is the actual average of the whole, since

$$£4{\cdot}15 \times 10{,}000 = £41{,}500,$$

the amount of the weekly Wage bill.

i. Definition of the Weighted Arithmetic Average.

This type of Average is known as the **Weighted Arithmetic Average**, and may be described as *an average whose component items have been multiplied by*

certain numbers known as "Weights" before being added, the aggregate being then divided by the sum of the weights, instead of by the number of items. If the numbers of men in each of the above groups of workers had been the same, then weights need not have been used, as the divisions being equal would have caused a Simple Arithmetic Average to suffice. This principle of the Weighted Average is very important in all cases where varying quantities are in evidence, and it will be necessary, for instance, to use it in a factory where average cost per unit of the commodities manufactured is desired, and the output varies for the various commodities, also where the average output per machine is required, and the machines are of different patterns or are working under varying conditions.

ii. THE WEIGHTS TO BE USED.

It is not, however, necessary to use the exact weights. They may be estimated, *but if the number of weights used is small, the size of those used may have a marked effect upon the result, and therefore the weights used should be approximately correct.* It has been demonstrated by Professor Bowley, in his book "The Elements of Statistics," that mathematically an *error in the Weights tends to be much less serious than a corresponding error in the size of the items,* for an error in the weights, if many of them are used, will in all probability be Unbiassed, and so tend to cancel or nullify one another. The following example will suffice to illustrate this fact.

Three countries have an acreage of 3,542,160, 2,673,940 and 1,340,920 respectively, under Corn crops, and produce an average of 35·4, 29·6 and 31·2 bushels per acre. What is the average yield per acre of the ground under cultivation ?

Case I. Using actual figures we get :—

$$\frac{(35\cdot4\times3{,}542{,}160)+(29\cdot6\times2{,}673{,}940)+(31\cdot2\times1{,}340{,}920)}{3{,}542{,}160+2{,}673{,}940+1{,}340{,}920}$$

$$=\frac{125{,}392{,}464+79{,}148{,}624+41{,}836{,}704}{3{,}542{,}160+2{,}673{,}940+1{,}340{,}920}=\frac{246{,}377{,}792}{7{,}557{,}020}.$$

$$= \mathbf{32\cdot6025\ bushels.}$$

Case II. If we now approximate the figures for the acreage to the millions, correct to the first decimal place, we get :—

$$\frac{(35\cdot4 \times 3\cdot5)+(29\cdot6 \times 2\cdot7)+(31\cdot2 \times 1\cdot3)}{3\cdot5 + 2\cdot7 + 1\cdot3}$$

$$=\frac{123\cdot9 + 79\cdot92 + 40\cdot56}{3\cdot5 + 2\cdot7 + 1\cdot3}=\frac{244\cdot38}{7\cdot5}$$

$$= \mathbf{32\cdot584\ bushels.}$$

a difference between the result obtained in *Case I.* of only ·0185 of a bushel.

Case III. If, however, we approximate the average yields to the nearest bushel, and work with the correct acreage, the result is as follows :—

$$\frac{(35 \times 3{,}542{,}160)+(30 \times 2{,}673{,}940)+(31 \times 1{,}340{,}920)}{3{,}542{,}160 + 2{,}673{,}940 + 1{,}340{,}920}$$

$$=\frac{123{,}975{,}600+80{,}218{,}200+41{,}568{,}520}{3{,}542{,}160+2{,}673{,}940+1{,}340{,}920}=\frac{245{,}762{,}320}{7{,}557{,}020}$$

$$= \mathbf{32\cdot52106\ bushels.}$$

a difference from the result obtained in *Case I.* of ·08144 of a bushel, a much larger error than is found in (*b*), and one which in this case would tend to grow as the numbers of acres involved were increased.

iii. When Weighted Averages must be Used.

(1) *When classes of the same Group contain a varying number of items.*

As already shown weights must be used in all cases when the number of examples falling into varying sub-divisions of the same grouping shows considerable divergence, and where it is desired to obtain an average which is representative of the group as a whole, for if weights be not used under such conditions observations and deductions will obviously be fallacious. Detailed investigation of the sub-divisions will, of course, always be more satisfactory, and then the Simple Arithmetic Average will usually suffice; but frequently it becomes necessary to find a unit for the whole which can be compared with others, and in such cases the Weighted Average must be used. If, for instance, we desire to compare the wages earned by the workmen in two factories engaged in producing similar articles and the following data is available, then the only average really representative of the factory will be the Weighted Arithmetic Average.

Description of Worker.	Factory A.		Factory B.	
	No. of Employees.	Weekly Wage per Employee.	No. of Employees.	Weekly Wage per Employee.
		£ s. d.		£ s. d.
Labourers	250	2 10 0	320	2 5 0
Apprentices	20	1 10 0	40	1 5 0
Semi-Skilled Workers	200	3 10 0	300	4 0 0
Skilled Workers ..	150	5 0 0	200	5 0 0
Total	620		860	

Supposing the Simple Arithmetic Average were used we find the results will be as follow :—

Factory A. $\frac{£2{\cdot}5+£1{\cdot}5+£3{\cdot}5+£5}{4} = \frac{£12{\cdot}5}{4} = \underline{\underline{\mathbf{£3{\cdot}125}}}$

or £3 2s. 6d. per week.

Factory B. $\frac{£2{\cdot}25+£1{\cdot}25+£4+£5}{4} = \frac{£12{\cdot}5}{4} = \underline{\underline{\mathbf{£3{\cdot}125}}}$

or £3 2s. 6d. per week.

Apparently then the average wage per worker in the two factories is the same; but we have, however, seen that such average when multiplied by the number of workers would not necessarily give the wages Bill of the factories. and until the average wage be correct comparison is impossible. The weighted averages will be as follow :—*

Factory A.

$$\frac{(£2{\cdot}5\times25)+(£1{\cdot}5\times2)+(£3{\cdot}5\times20)+(£5\times15)}{25+2+20+15}$$

$$= \frac{£62{\cdot}5+£3+£70+£75}{62} = \frac{£210{\cdot}5}{62} = \underline{\underline{£3{\cdot}39516}}$$

= £3 7s. 11d. per week.

Factory B.

$$\frac{(£2{\cdot}25\times32)+(£1{\cdot}25\times4)+(£4\times30)+(£5\times20)}{32+4+30+20}$$

$$= \frac{£72+£5+£120+£100}{86} = \frac{£297}{86} = \underline{\underline{£3{\cdot}45348}}$$

= £3 9s. 1d. per week.

* It will be observed that the weights have been approximated by leaving out the last cipher. They might, without affecting the result, have been approximated to hundreds, correct to the first decimal place.

A marked difference in average wages is now displayed. Too much stress cannot however be placed upon the fact that while this average is that for the whole of the workers, it will not really represent the actual conditions of any one class, or any individual worker, but is useful only for the purposes of a general comparison, and is useful only in surveys of an industry as a whole. To study the problem properly we should compare the classes separately, and so obtain reliable information as to the actual position of the worker.

(2) *When Proportions Change.*

In the same way, if Wages are increased and the numbers of the workers in the different classes also change, the alteration in the average would not necessarily be that anticipated. If we again refer to the case above, of 10,000 workmen (see page 83), and assume that all the workers, irrespective of the grade to which they belong, received an advance of £1 per week, the average wage would then become :—*

$$\frac{(£6\times 40) + (£5\times 35) + (£4\times 25)}{40+35+25}$$

$$= \frac{£240 + 175 + 100}{40 + 35 + 25} = \frac{£515}{100} = \underline{\mathbf{£5{\cdot}15}}$$

$$= \mathbf{£5\ 3s.}$$

If at the same time as the change in wages was made the number of skilled men employed was reduced by 1,000, the semi-skilled by 500, while the unskilled workers were increased by 1,500, the number of workers would be unchanged, but the average wage would now be :—*

* Weights have been taken in hundreds by dropping the last two ciphers, thus reducing the arithmetical work involved. In the second case the same result would have been obtained by using thousands, *i.e.*, dropping the last three ciphers.

$$\frac{(£6 \times 30) + (£5 \times 30) + (£4 \times 40)}{30+30+40}$$

$$= \frac{£180 + 150 + 160}{30 + 30 + 40} = \frac{£490}{100} = \underline{£4{\cdot}9}$$

$$= £4\ 18s.$$

an average increase over the rate prior to the change of 4·9 — 4·15 = £0·75, or only 15s. per week. It can be taken as an axiom, therefore, that if the number of men in any class changes, or the proportion in each class alters, the weighted average will change also.

iv. Commercial Uses of the Weighted Arithmetic Average.

Generally speaking the Weighted Arithmetic Average enjoys the same advantages, suffers from the same disadvantages, and has the same uses as the Simple Arithmetic Average. The Moving and Progressive Averages can, of course, be applied to the Weighted Arithmetic Average with equal facility and usefulness as to the Simple Arithmetic Average, and the principle is frequently used by merchants to ensure that their dealings will yield them an ultimate profit, though individual transactions may apparently show a loss. The following series of transactions will illustrate the method :—

TABLE G.

Table illustrating the Application of the Progressive Average.

Date.	No. of Bales Purchased.	Price per Bale.	Amount Paid.	Total Bales Purchased.	Total Amount Paid.	Average Price per Bale.
		£	£		£	£
January 1	100	2·0	200	100	200	2·0
	100	1·8	180	200	380	1·9
	50	1·4	70	250	450	1·8
January 2	250	1·6	400	500	850	1·7
	100	2·3	230	600	1,080	1·8
January 3	400	1·3	520	1,000	1,600	1·6

As each purchase is made a Weighted Average is extended showing the average cost per unit of the goods purchased, and from this progressive average the merchant can calculate the price at which he can sell the goods, or part of them, to ensure a profit. In the foregoing example, after the merchant had made his first purchase, he could not sell any of the goods under £2 per bale, unless he made a loss, but after the second purchase had been made he could afford to sell the whole of his 200 bales at £2 per bale, and then make a profit, while his third purchase brought the average cost per bale down to £1·8, and therefore a sale made at anything over this figure would yield him a profit. This system would have to be adopted for every commodity separately, and is particularly suitable when dealing on a fluctuating market. The same procedure is also frequently adopted on the Stock Exchange by Speculators who desire to snatch a profit where a loss seems imminent. Supposing a purchase is made of 1,000 shares at 30s. per share in anticipation of the price rising and being able to sell at a profit, but the anticipation is not realised, and instead of advancing, the price fell steadily, thus not only is a profit apparently impossible, but a loss seems inevitable. Let us assume that the shares fell to 25s., at which price a further purchase of 2,000 shares is made. The average cost per share of the 3,000 shares would now be :—

$$\frac{(£1{\cdot}5 \times 1{,}000) + (£1{\cdot}25 \times 2{,}000)}{3{,}000} = \frac{£1{,}500 + 2{,}500}{3{,}000}$$
$$= £1\ 6s.\ 8d.$$

so that if the price subsequently advanced to 29s. per share, they could be sold to yield a profit of 2s. 4d.

per share, or a total profit on the whole of the transactions of 2s. 4d. + 3,000 .. £350, although the Selling Price did not reach the figure at which the first 1,000 Shares were bought.

Such a Progressive Weighted Average would enable many a Trader to increase his business by adjusting his prices to market conditions. It might be argued that the compilation of such records takes time, but on the other hand most business men use the system in some form or another, and the only additional thing, therefore, is to do it systematically, instead of, as is the general rule now, when the market is moving rapidly against the operator. If it were always available, transactions might be carried out successfully in the face of severe competition, when competitors had not the same data available.

(c) The Descriptive Average.

If the data on which we are working are not exact, or are incomplete, it follows that the averages compiled from such data will contain an unknown and immeasurable error. This happens when 100 Workers are described as each earning between £3 10s. and £4 per week, and we assume the average wage to be £3 15s. which may be entirely incorrect, as we have estimated the numbers receiving varying rates of pay as being equal, whereas they may vary considerably. Such an average may be described as a "**Descriptive Average,**" and it is then known that it cannot be relied upon for further working. It, however, sometimes becomes advisable to obtain a descriptive average as a guide—but as a guide only—to varying conditions, and the method of obtaining it is illustrated in the following example :—

What is the "average" wage of 75 workers, of whom 10 earn 10s. or less per week, 20 earn over 10s. and not over 20s., 30 earn wages exceeding 20s. and up to 30s., and the remaining 15 earn over 30s., but not over 40s. ?

It is necessary to assume that the wages are evenly distributed over the groups (*i.e.*, in the first group the first worker receives 1s., the second 2s. and so on), and the average rate received by the men in the first group of ten men would therefore be:— $\frac{1+2+\ldots+10}{10} = 5.5$s., and we assume that each group receives the same average plus the number of units of the difference of the wages between the groups (in this case 10s.), by which the wages of the succeeding groups exceed the first group. When this is done we get :—

Group No.	(*a*) Rate of Wages	(*b*) No of men in Group	(*c*) Wages.	(*d*) Wages in Units of 10s.	(*e*) Product of columns (*b*) and (*d*).
			s.		
1	10s. and less ..	10	say 5 5	5·5 + 0	0
2	Over 10s. and not exceeding 20s ..	20	5 5 + 10	+ 1	20
3	Over 20s. and not exceeding 30s...	30	5·5 + 20	+ 2	60
4	Over 30s. and not exceeding 40s..	15	5 5 + 30	+ 3	45
		75			125

Columns (*b*) and (*e*) are then totalled, and the following is the formula to find the average :—

Assumed average of first group

$$+ \left(\frac{\text{Total of column } (e)}{\text{Total of column } (b)} \text{ of the unit of increase}\right)$$

so that in the above case we get :—

$$5{\cdot}5 + \left(\frac{125}{75} \text{ of } 10\text{s.}\right) = 5{\cdot}5\text{s.} + 16{\cdot}6\text{s.} = \mathbf{22s.\ 2d.}$$

If the numbers in column (*b*) are large, then we can use approximate numbers in place of the actual numbers without materially affecting the result. Thus in the example given above the numbers in column (*b*) might have been approximated to 1, 2, 3 and 1·5 respectively.

It will thus be seen that the Descriptive Average is a combination of the Simple and the Weighted Arithmetic Average. The Simple Arithmetic Average is used to obtain the basic rate for the first group, and the Weighted Arithmetic Average for obtaining the amount to be added to this basic rate to obtain the Descriptive Average. It will be observed also that the whole of the data is of an indefinite nature and that the assumption has to be made in obtaining the basic rate that the workers are paid, not by a flat rate but in an arithmetic progression. Obviously, therefore, the result cannot be relied upon, but still is useful for comparing results when the data in each or all cases are equally indefinite.

§ 3.—The Geometric Average.

There is also the "**Geometric Mean,**" or Average, which is found in the following manner : If the number of items in a series = n, we multiply the items together

and extract the nth root of the product. It is difficult of calculation, and often involves much work even when the aid of Logarithms is invoked. This average, however, possesses one great advantage over the Arithmetic Average in that it tends to nullify the effects which items of very large magnitude have upon the Average, and this is important when such extremes are few in number as compared with the remainder of the data. If, for instance, it be desired to ascertain the average of the magnitudes, 2, 3, 4, 5, 6, 7, 8, 9, 10 and 181 we find the Arithmetic Average would be

$$\frac{2+3+4+5+6+7+8+9+10+181}{10} = 23{\cdot}5$$

a result which cannot be said to be in any way representative of the magnitudes involved, for it has, as it were, been pulled away from the bulk of the data by the one very high magnitude at the upper extremity. The Geometric Average would give us the following :—

$$\sqrt[10]{2\times3\times4\times5\times6\times7\times8\times9\times10\times181} = 7{\cdot}6164$$

thus bringing the average within the range of the majority of the examples in the data, and so making it more representative. The Geometric Mean is always lower than the Arithmetic Average.

Owing, however, to the difficulties of calculation, and the fact that unless some of the extremities are of considerable magnitude, the results vary but little from those obtained from the use of the Arithmetic Average,* it will not be used in Commercial work to any extent.

* This fact will be found demonstrated in the Table on page 251.

§ 4.—Median.

(a) Arraying the Data.

If a number of similar objects are placed side by side in order of their magnitude (*i.e.*, the smallest on the left and the largest on the right, or *vice versa*), they are said to be "**Arrayed,**" and the middle example of the array is known as the "**Median,**" and may be said to be that item whose magnitude most closely corresponds to the magnitude of all the other examples in the array, when such examples are numerous. Diagram No. 1 (see page 142) shows eighty items, ranging in size from 1 to 70, arrayed, and the sizes may represent any data desired. The centre item is the Median. In such a diagram the size of any item is easily ascertained by reference to the vertical scale. This example is a very simple form, but frequently our information is not so definite, and our examples have to be grouped as falling within certain limits, so that to plot these accurately is impossible, as the size of the items in the different groups is only known to be within certain limits, and hence the magnitude of the Median cannot be accurately determined. When the number of examples is an even number the Median does not actually exist, but is deemed to be fixed midway between the two middle items, so that in the above example the Median is half way between the fortieth and forty-first items.

(b) Frequency Tables.

Indefinite data is illustrated in the following case. There is a group of 261 workers, and of these 50 earn not more than 10s. per week, 70 earn over 10s. but not exceeding 12s., 60 earn from 12s. to 14s., and 81 take from 14s. to 16s., and it is desired to ascertain what is

the probable wage of the Median example. We first compile what is known as a "Frequency Table," which is constructed as follows:—

Group No.	(a) Amount of Wages (1)	(b) Number of Workers receiving such Rate of Wages. (2)	(c). Cumulative Numbers. (3)
1	Not exceeding 10s.	50	50
2	Over 10s., but not exceeding 12s.	70	120
3	Over 12s., but not exceeding 14s.	60	180
4	Over 14s., but not exceeding 16s.	81	261

(1) Column (*a*) may be described as the "Size" of the data under consideration.

(2) Column (*b*) gives the "Frequency," *i.e.*, the number of the particular size under review.

(3) When Columns (*a*) and (*b*) only are given, the table is known as a Frequency Table, but when column (*c*) is added it becomes a Cumulative Frequency Table.

From the data given in columns (*a*) and (*b*) we cannot obtain the Median by inspection, and hence it is necessary to re-group the classes by taking the Cumulative Frequency as shown in column (*c*), from which we obtain the information that there are 120 workers earning varying sums of not more than 12s. per week, 180 take anything not exceeding 14s. per week, and so on. The Median is the 131st case, *i.e.*, $\frac{n+1}{2} = \frac{261+1}{2} = 131$; and will now be seen to be in group No. 3, *i.e.*, the group in which the workers earn sums varying from 12s. to 14s. per week, and is the eleventh case of that group. We therefore know that the particular individual earns between 12s. and 14s. per week, and if it be desired to ascertain what is the approximate amount of his earnings it is necessary to assume that the "CLASS INTERVAL" (*i.e.*, the difference between the highest wages in the class below the one in question and the

lowest of the class above) is spread evenly over the workers in the group. This is known as " **Interpolation.**"

(c) Rule for Calculating the Magnitude of the Median.

The earnings of the Median man may now be estimated to be :—

The highest amount received in the group below

$$+\left(\frac{\text{Median number—Total number of cases to end of class below}}{\text{Number of men in the class in which the Median is situated}} \text{ of the class interval}\right)$$

Apply this to the data we get :—

$$12\text{s.} + \left(\frac{131-120}{60} \text{ of } 2\text{s.}\right) = 12\text{s.} + \left(\frac{11}{60} \text{ of } 2\text{s.}\right)$$

$$= 12\text{s.} + 4{\cdot}4\text{d.} = \mathbf{12s.\ 4{\cdot}4d.}$$

A simpler formula than that given above would be :—

The highest amount received in the group below

$$+\left(\frac{\text{Position of Median in group containing it}}{\text{Number of men in group containing the Median}} \text{ of the class interval}\right)$$

$$i.e., \ 12\text{s.} + \left(\frac{11}{60} \text{ of } 2\text{s.}\right)$$

which gives the same result as above.

(d) Quartiles.

If we decide to divide the array further we can take the items placed midway between the Median and the extremes, *i.e.*, we divide the array into four parts. The units which so divide the array into four equal parts are known as "**Quartiles.**" In the Diagram No. 1

(page 142) the items 20 and 60 will be the Quartiles, the former, which is between the lower extreme and the Median, is known as the "**Lower Quartile**," while the latter being between the Median and the upper extreme is called the "**Upper Quartile.**" Like the Median, the Quartiles are merely representative. They are easily ascertained from a graph, and their position can usually be ascertained by calculation, $\left(i.e., \text{ The Lower Quartile} = \frac{n+1}{4}\right.$; and the Upper Quartile $= \left.\frac{3\,(n+1)}{4}\right)$; but when our data is in the form of a frequency table it is better to locate them in the same manner as the Median. In the foregoing Frequency Table the Lower Quartile is the 65·5th case, and is between the fifteenth and the sixteenth items in the second group, which contains seventy cases. Its magnitude therefore is :—

$$10\text{s.} + \left(\frac{15{\cdot}5}{70} \text{ of } 2\text{s.}\right) = 10\text{s.} + 5{\cdot}314\text{d.} = \mathbf{10s.\ 5{\cdot}314d.}$$

(e) Deciles and Percentiles.

An array may also be divided into Tenths, in which case each of the tenth items is known as a "*Decile*," or into hundredths, in which case the representative items are called "*Percentiles*."

(f) Advantages of the Median.

i. It is usually an actual example from the data under review, and so is often more representative than the Arithmetic Average.

ii. It is capable of exact location.

iii. It can be obtained by inspection.

iv. It is not affected to any material degree by the magnitude of the extremities of the data under review.

v. It can be ascertained without knowing the magnitude of the extremities, provided the number of cases is known.

vi. It is capable of being physically measured.

vii. It can be obtained with reasonable exactitude when data falls into groups or classes, the limits of which can be defined, though the items in such groups or classes are incapable of exact measurement.

(g) Disadvantages of the Median.

i. It is incapable of measurement by any simple mathematical formula.

ii. The data must be arrayed before it can be located, an operation which often entails considerable work.

iii. The aggregate of the items under review cannot be obtained after it has been measured and the total number known, as is the case with the Arithmetic Average.

iv. It cannot be used for further work involving mathematical calculation.

v. If the items in the array vary very greatly in magnitude it may not be truly representative.

(h) Uses of the Median.

The Median is of great utility for comparing data which it is impossible or difficult to measure individually, and which have to be grouped within certain

defined limits. It is particularly useful therefore when considering social phenomena or conditions of all kinds such as Wages (from an Abstract point of view); the Distribution of Wealth, Ability, etc. From a Commercial point of view, however, it will probably be considered too indefinite and too isolated to be used to any great extent, for the magnitude of Commercial data will usually be very widely dispersed owing to the difference which exists in the element of management, and hence the Median may not be truly representative.

§ 5.—The Mode.

In data of all kinds, and particularly in those relating to natural phenomena, it is found that some particular type or magnitude is constantly recurring, so much so that after proper investigation has been made it can be said what will be the most likely type or magnitude to be met with when further samples are taken. In other words, if the data under review be arrayed in order of magnitude, there will be found a predominating group of the same, or approximately the same magnitude. This is illustrated in Diagram No. 1 (page 142) where it will be seen that of the eighty items therein arrayed there are ten (*i.e.*, one-eighth of the whole) which are uniform in magnitude, while the other seventy items vary in size from 1 to 70. The items in question are those numbered from 40 to 49 and are of size 40. If we were to shuffle the whole of the items comprised in this array, and pick out an item at random, it is obvious that we shall have ten times more chance of securing one of the size 40, than we shall of drawing an example of any other size. Such predominant items are known as the "**Mode**" or

"**Norm,**" which may therefore be defined as the predominant kind, type or size of item, or the position of greatest density. When we speak of the "Average Man" we really mean the type of man who is most likely to be met with, seeing that we cannot convert ability, personality or character into arithmetic terms, and it would therefore be better in many ways to refer to him as the "Modal Man," rather than as the Average Man.

(a) Method of Locating the Mode.

When items are arrayed, or when data is presented in the form of a diagram as in Diagram No. 1, or in the form of a curve or graph as in Diagram No. 2 (page 144), the position of the Mode is easily distinguishable from the fact that when numerous examples of one magnitude are present the array "flattens out," a process clearly seen in Diagram No. 2. It frequently happens, however, that magnitudes vary but little, and that the number of examples of several magnitudes may be almost equal, and in such cases it is more difficult to locate the real mode by inspection. It then frequently becomes necessary to split the array into groups containing various magnitudes, and continue the process until one group stands out preeminently. This enables the Modal group to be identified, though even then the Mode itself can often only be approximately located, though the limits or error are usually very small.

The following Table shows the method of regrouping. Column (*a*) shows the size of the items, and (*b*) the frequency, or the number of items of each particular size.

Size	Frequency.					
(a)	(b)	(c)	(d)	(e)	(f)	(g)
10	34	72				
11	38		80	114		
12	42	88			126	
13	46		94			136
14	48	99		145		
15	51		101		149	
16	50	93				144
17	43		94	144		
18	51	100			143	
19	49		91			142
20	42	80		129		
21	38		74		116	
22	36					

It will be seen from Column (*b*) that no item has such a frequency that it stands out pre-eminently, though the item of size 15 or 18 with 51 examples would appear for the moment to be the Mode, but as the frequencies of many of the items are so similar it is necessary to ascertain more definitely if our first inspection is correct. We shall therefore group the items into fresh classes with a view to ascertaining if any other factor is likely to show a greater predominance. Column (*c*) shows the number of examples in the new groups, *e.g.*, it will be seen that there are seventy-two examples in the group composed of those items of sizes 10 and 11, and so forth. The process of grouping the items in pairs is continued in column (*d*), but sizes 11 and 12,

13 and 14, and so on would be combined, thus each of the original classes is combined with that before and after it. In column (*c*) the Mode would appear to be in that group which contains items of the sizes of 18 and 19, as there are 100 examples in this new group, but in column (*d*) the largest group is seen to be that of sizes 15 and 16, and as the location is still no clearer it is necessary to regroup the data. This time we combine three of the original classes, starting with the first three. The result is seen in column (*e*). Column (*f*) is then prepared, again grouping in classes of three, and this is done by shifting the limit of the group one class down, and the process is repeated in constructing column (*g*). When we group two classes together we do so twice, moving the upper limit of the group down one class the second time, and when grouping three classes we do so three times in all, proceeding in a similar manner by shifting the upper limit down one each time. In this way we group each item with those preceding and succeeding it in such a way that every combination of grouping possible is used. Supposing we desire to group the item 5 with as many groups as possible we get the following combinations :—

Group including

Two items.	Three items.	Four items.	Five items.
4, 5.	3, 4, 5.	2, 3, 4, 5.	1, 2, 3, 4, 5.
5. 6.	4, 5. 6.	3, 4, 5, 6.	2, 3, 4, 5, 6.
	5, 6, 7.	4, 5, 6, 7.	3, 4, 5, 6, 7.
		5, 6, 7, 8.	4, 5, 6, 7, 8.
			5, 6, 7, 8, 9.
No. of Groups possible 2.	3.	4.	5.

If no group is predominant after grouping in threes it will be necessary to group four classes together four times, and so on till we get a predominant group, though it must be borne in mind that each time we find it necessary to bring another class into the new group we are making the location of the Mode more approximate. In the example under review we find that columns (*f*) and (*g*) show groups which are distinctly larger than the others, and moreover, the largest group seems to centre on the same spot. An analysis will show which class will appear constantly in the maximum group. This analysis can be prepared as follows :—

Items comprising maximum groups in each column shown :—

Col. (*b*)			15;			18;	
(*c*)						18;	19;
(*d*)			15;	16;			
(*e*)	13:	14:	15;				
(*f*)		14;	15;	16;			
(*g*)			15;	16;	17;		

It will at once be seen from such an analysis that the item of size 15 appears in the maximum group five times in six cases, and we therefore assume that 15 is the Mode of the items under review, a fact which is confirmed if further grouping be carried out. In other words if other examples be collected we shall expect size 15 to be met with more often than any other magnitude. Quite frequently we cannot locate the Mode so definitely as two sizes may constantly recur in the maximum groups, and in such cases, while the exact Mode cannot be ascertained, it is possible to define the limits within which it will occur.

The above procedure is applicable when the data is in the form of a Discrete Series, but in a Continuous Series it only locates the class which contains the Mode.

It is sometimes desirable to ascertain the location of the Mode within the class, and this is done from the following formula :—

$$Z = l + \frac{f_2 c}{f_2 + f_1}$$

where :

Z = the Mode.

l = the lower limit of class.

c = the class interval.

f_1 = the number of items in the next lower class.

f_2 = the number of items in the next higher class.

Example:

TABLE SHOWING FREQUENCY WITH WHICH PROFITS ARE MADE.

Profits.	Frequency.
Exceeding £3,000 and not exceeding £4,000	3
,, £4,000 ,, ,, ,, ,, £5,000	7
,, £5,000 ,, ,, ,, ,, £6,000	22
,, £6,000 ,, ,, ,, ,, £7,000	60
,, £7,000 ,, ,, ,, ,, £8,000	85
,, £8,000 ,, ,, ,, ,, £9,000	32
,, £9,000 ,, ,, ,, ,, £10,000	9
	n = 218

The Modal Group is obviously £7,000 to £8,000, but to find the actual Mode we apply the formula given above and thus get :—

$$Z = £7{,}000 + \frac{32 \times 1{,}000}{32 + 60}$$

$$= £7{,}000 + \frac{32{,}000}{92}$$

$$= \underline{\underline{\mathbf{£7{,}347{\cdot}826}}} = \text{the required Mode.}$$

(b) Advantages of the Mode.

i. Like the Median it can often be obtained by inspection, an advantage which the Arithmetic Average does not enjoy.

ii. Except when the data forms a Continuous Series it is not only an actual case, as is the Median, but it has the great advantage over the Median that it is not an isolated example, *but is the predominating magnitude, and therefore is the example which is most likely to occur if further examples be collected.*

iii. The extremes, however great their magnitude, do not affect the Mode at all, unless they are very numerous, and thus extreme variations are eliminated.

iv. It can be ascertained without knowing what the extremes are, always provided it be known that the usual natural law relating to extremes is operating. Abnormalities in magnitude are, however, comparative rareties, and can usually be ignored.

v. Like the Median it can be ascertained when data falls into groups or classes, the limits of which can be defined, though the items in such groups or classes are incapable of exact measurement.

vi. It is clearly understood by the ordinary man.

(c) Disadvantages of the Mode.

i. It is incapable of being ascertained by mathematical formulæ.

ii. Like the Median it cannot be measured by any simple arithmetic process.

iii. The data must be arrayed in similar manner to that required to obtain the Median.

iv. No aggregate can be obtained even when the magnitude of the Mode and the number of items in the array be known, and consequently it cannot be used if further investigation entailing mathematical calculations is to be made.

v. A comparatively small number of items of the same magnitude may determine the Mode when, however large the number of examples under review may be, they are scattered over a wide range of sizes.

(d) Commercial Uses of the Mode.

The Mode has so far not been used to any extent in Commerce, probably because its functions and characteristics have not been thoroughly understood, but since it is the outstanding feature of reviewed data, and if this data has been correctly collected the most representative item is that which is most likely to occur if further examples of the same nature be collected. It should thus prove a most reliable guide to possible results when similar work is to be done, or identical conditions prevail. Meteorological forecasts are really based on the use of the Mode, and such forecasts are daily becoming more important to the community.

In studying Output the Mode should prove of great advantage. By recording the Output of a number of similar machines engaged on the same class of work, it will be found that though the Output per machine varies, there will be a certain number of machines giving more or less the same Output over a period, thus

establishing a Modal Output per machine. Such Modal Output will form the basis of comparison, and will direct attention to the varying skill of the workers, or the efficient working of the machines. In one case, known to the writer, a factory engaged on repetition work increased output, and decreased the cost per unit of output to a considerable degree by establishing a Modal standard per machine, and then carefully studying the reasons why other machines did not achieve similar results. This led to faults in the machines, methods of transmission, and lack of skill or care on the part of operatives being revealed, and incidently, by encouraging team work, reduced idle time through breakdowns very materially, as well as the cost of repairs, since operatives immediately reported anything which tended to reduce their output, and defects which might have developed were remedied before they had an opportunity of becoming serious. A little extra work was naturally involved in recording the daily output per machine, but this was more than counter-balanced by the increased efficiency and output. Such records also made it possible to estimate with a greater degree of accuracy the length of time necessary to complete certain work, or obtain a certain output.

Similarly, if the time taken to perform certain operations be recorded, a Modal Time will probably be revealed, and this will, of course, represent the most likely time which will be taken to perform similar work under similar conditions. This information would prove of great assistance when entering into a contract entailing a time element, and would also materially aid the estimator in fixing the probable cost of performing certain work, or producing certain

goods, and thus lead to better business, more reliable records, and larger profits, seeing that the business man with the greatest knowledge and experience is always bound to do better than his more poorly equipped competitor.

PART II.

PRESENTATION AND USES OF STATISTICAL DATA.

SYNOPSIS OF CHAPTER V.

TABULATION.

§ 1. Importance of Proper Tabulation.

2. Kinds of Tabulation.
   (*a*) Simple.
   (*b*) Complex.

3. Rules for Tabulation.

4. Data to be Tabulated.

5. Title of the Table.

6. The Uses of Percentages and Averages in Tabular Work.

7. Tabulation of Commercial Data.

CHAPTER V.

TABULATION.

However carefully our records are compiled, it is essential that they be presented to those for whom they are intended in such a manner as will enable them to make the best use of the information given, and, moreover, so that they will at once realise the importance of significant factors which appear in the data under review. The method of presentment is important, since a good cause may be irreparably damaged if the case be not correctly placed before the people interested, while a poor case may win support if the facts are carefully, logically and succinctly drawn up so as to excite the interest of the reader. It has been said that the facts to be presented do not matter so much as the method of presenting them, and this statement is in some measure true in everyday life, for the speed of the World's commerce as conducted to-day is such that business opportunities with their resultant profit may be lost if time has to be spent in an exhaustive study of data, which may or may not assist the business man in coming to a decision. The important data of a business must be available in such a form that they can be speedily obtained, easily understood, and their

significance grasped without undue study and waste of time. Facts must therefore be presented in such a way that they attract the attention, give the information required, and enable the results following on a definite action to be foreseen and provided for. The presentment of data has now reached a high standard in some of our Newspapers, the headlines of which supply such information that the busy man is enabled to grasp the salient features of the news without the necessity of voluminous reading of details, which frequently are of minor importance. Time is thus saved for concentration on the affairs which are of more vital interest.

§ 1.—Importance of Proper Tabulation.

In Statistical work we rely upon two forms of presentation, viz.: Tabulation of Results and Graphic Presentment, and both present some difficulties in their preparation. Even when the Graphic method is used Tabulation is necessary, though sometimes only to a limited extent, before the form the diagram is to take can be decided upon. The Tabulation of results is really the final stage of the work of compilation, and subsequently becomes the basis for the consideration, deduction and application to present and future problems. It consists of the classification of the data obtained by the inquiry, sampling, recording and averaging. In preparing our Tables we must carefully consider the reason for which the data are required or have been assembled, and also the ultimate uses to which they will be put, and we can then decide the form the Tabulation shall take, in order that the information may be shown clearly, succinctly and yet briefly. Tabulations must be

capable of being easily read, and must necessarily, therefore, be compact in form, and should be so arranged that those figures which have to be compared will be in close proximity to one another. This has already been demonstrated in Table F (page 79), where the information contained in column (*c*) was placed before that in column (*d*), although the latter data had to be obtained before the details given in the former column could be calculated. A small Table enables all these points to be dealt with easily, but frequently our data are of such magnitude that further divisions become necessary, and when this is the case the Table tends to become unwieldy, and its value is decreased in an inverse ratio to its size. Generally speaking, each Table should be definitely prepared to show the particular point it is desired to demonstrate, for planning a tabular statement to enable comparison to be made between several distinct and differing units is apt to be confusing to those not well acquainted with the particular details under review. Care must be also taken to see that the units which are being compared are such as will enable comparison to be made. It would, for instance, serve no very useful purpose to show in the same statement, with the object of comparison, the *Output* of a business and the *Gross Profit*, since Gross Profit bears no direct relationship to the quantity of goods produced; but, on the other hand, *Sales* (or Turnover) can be compared with *Gross Profit*, since the latter depends to a large measure upon the former, though it must be borne in mind that profits will not of necessity fluctuate to the same degree as the Turnover, for even if the price of the commodity rises, but expenses change in a different ratio, the profit will not increase to the same extent as the Turnover.

§ 2.—Kinds of Tabulation.

We can prepare our Tabular Statement in either a **Simple** or a **Complex** form.

(a) Simple.

Simple Tabulation is designed to supply information in respect of one or more groups of independent inquiries. The following Table is an example of Simple Tabulation :—

TABLE H. *

NUMBER AND MEMBERSHIP OF SOCIETIES ENGAGED IN THE CO-OPERATIVE MARKETING OF MEMBERS' PRODUCE, IN EACH OF ELEVEN YEARS :—

Years.	Number of Societies.	Membership
1	124	9,732
2	129	10,401
3	140	11,943
4	159	13,701
5	228	25,384
6	289	40,697
7	315	45,779
8	302	44,201
9	263	40,309
10	248	38,395
11	230	38,060

(b) Complex.

Complex Tabulation shows the division of a Total into two or more categories, as in Table J (opposite), where the total number of failures recorded in the United Kingdom are divided into those for the Wholesale Trade, Professions and Retail Trades respectively.

Another form of Complex Table is one in which the separate columns can be totalled, and the grand total checked by the cross-casts of the columnar table, of which Table K (p. 119) is an example.

* Compiled from the "Ministry of Labour Gazette."

TABLE J.

FAILURES RECORDED IN THE UNITED KINGDOM DURING NINE YEARS.

Year.	Wholesale Trades.	Professions.	Retail Trades.	Total Failures.
1	649	668	5,440	6,757
2	130	268	608	1,006
3	310	356	1,598	2,264
4	1,028	549	3,971	5,548
5	1,246	717	5,627	7,590
6	1,224	739	6,278	8,241
7	1,081	783	6,270	8,134
8	962	774	6,242	7,978
9	940	671	5,639	7,250

§ 3.—Rules for Tabulation.

1. Extra Columns may be added, or additional information tabulated, at the discretion of the compiler, but it must be remembered that the more minute the divisions the greater the accuracy of the Table becomes. On the other hand, every additional column makes the Table more difficult to read, causes comparison to be less effective, and is likely to lead to confusion of thought and possible error, all of which are serious disadvantages, particularly to the busy commercial man.

2. The convenience of the person who uses the Table must be consulted.

3. It must be so constructed as to be easily read, easily understood, and its figures easily compared and followed without unnecessary waste either of time or thought.

4. Tables which are too large, and therefore unwieldy should be split into sections, and these drawn together into a concrete whole by means of a summarised table.

5. No Table should be of such a size that the eye finds it difficult to take in the whole at a glance.

6. When deciding upon a number of columns it is important to place figures, which are intended to be compared, as near to each other as is practicable, and preferably in vertical formation rather than horizontal, for the eye has greater difficulty in traversing a horizontal line of figures than a vertical column.

7. Where actual figures *and* percentages or averages appear in the same Table, the percentages or averages should be placed in close proximity to the figures upon which they are based.

TABLE K *

ENGINE HOURS "IN TRAFFIC" (INCLUDING DEPARTMENTAL) OF THE COMPANIES' ENGINES FOR THE MONTH OF JUNE.

Railway Company.	Engine Hours.							
	Train.		Shunting.		"Other" excluding Departmental.		Depart-mental.	Total
	Coaching.	Freight.	Coaching	Freight	Coaching.	Freight	--	
Great Western	**158,241**	**127,371**	**35,006**	**214,995**	**16,327**	**26,787**	**21,811**	**600,538**
London & North Eastern ..	**254,890**	**231,853**	**51,454**	**359,156**	**28,746**	**34,522**	**62,813**	**1,023,434**
Southern Area	153,809	140,654	31,265	185,044	19,329	21,442	40,156	591,699
North Eastern Area	63,113	49,084	13,175	105,760	4,129	5,567	15,883	256,711
Scottish Area..	37,968	42,115	7,014	68,352	5,288	7,513	6,774	175,024
London Midland & Scottish	**324,111**	**310,826**	**83,968**	**411,339**	**33,173**	**53,642**	**70,936**	**1,287,995**
Western Area	171,985	156,994	49,014	218,942	17,028	34,270	40.174	688,407
Midland Area	96,927	115,779	21,723	133,944	7,350	13,652	22,730	412,105
Northern Area	55,199	38,053	13,231	58,453	8,795	5,720	8,032	187,483
Southern	**185.653**	**42,717**	**39,565**	**94,389**	**21,044**	**9,048**	**19,808**	**412,224**
Cheshire Lines Committee*..	8,030	9,569	1,744	11,000	410	886	1,230	32,869
Metropolitan	30,170	1,492	1,409	1,647	338	129	740	35,925
Midland & G.N. Joint ..	3,415	3,954	121	3,580	59	147	430	11,706
Great Britain	**1,038,121**	**724,786**	**212,387**	**1,102,178**	**100,061**	**125,459**	**178,999**	**3,481,991**

* Compiled from the Ministry of Transport's Monthly "Railway Statistics."

8. If Totals are the principal figures to be compared, they are best placed at the head of a vertical formation, or at the left-hand side of horizontal tabulation, but the reverse procedure is more advantageous when the subsidiary columns are those of primary importance, as is shown in Table K (page 119). No hard-and-fast rules on this matter can be laid down. Any departure from the practice adopted is apt to result in confusion to the reader, and hence affect the value of the Table adversely.

9. If it should happen that some of the details required to complete the Table are missing or cannot be obtained, then the total of the column from which such omissions have been made should either not be inserted, or shown in a different type, in order to draw attention to the fact that such total cannot be used for the purposes of comparison with other totals which are the summation of the whole of the data of the period or column to which it relates.

10. In deciding the number of columns into which the matter is to be tabulated due regard must be given to the relative importance of the different data, and the ruling of the Table should closely follow this relative importance. The principal sections should be divided by heavy or multiple rulings, and the breadth of the ruling should vary with the importance of the sub-divisions.

11. It will frequently be found that some of the data do not fall within the divisions decided upon, and in such cases a "Miscellaneous" column becomes a necessity, but care must be taken to see that these examples are not of a widely varying nature, as otherwise footnotes and references will be necessary, and if this happens the Table is not sufficiently self-

contained to give the best results. When the Miscellaneous items vary in character and are numerous, it is better to make a subsidiary table for all such items, and tabulate them in such a manner as will obviate the necessity for notes and references to explain them.

12. The divisions into which the matter is divided should be of such a nature that the data are adequately distributed over them, for a Table showing some columns full, and others with numerous blank spaces is apt to be confusing to the reader, therefore divisions which contain regularly recurring data should be placed together, so that others where the data is irregular in its occurrence do not intervene.

13. Care must also be taken to "range" the columns, *i.e.*, place millions under millions, thousands under thousands, and so on. This makes the Table more readable, and aids comparison and computation.

14. A number of lengthy columns of numerical data containing a large number of digits in each individual number should, whenever possible, be avoided, and thus approximation should be used, provided that the data does not lose any of its value in so doing; for example, it would be unwise to approximate the figures when the Tables are intended to be used for reference rather than for drawing reasoned conclusions therefrom.

§ 4.—Data to be Tabulated.

Having settled the object for which the Table is to be compiled, the next point to be considered is whether we shall tabulate the actual figures, or express them as ratios or averages, and this can only be decided by a careful consideration of the data and the object for which they are being used. The Output

of a Factory, the Turnover of a Business, the Imports or the Exports of Goods, the Goods carried by Vans or Wagons, are all examples where the actual figures would be used in tabulation. The Profits of a Company, the Dividends paid to Shareholders, the Number of Hours engaged on Productive Work would always be better expressed in the form of percentages or ratios; while the Weight of Goods carried per Van or Wagon, the Output of a Factory for each unit of time are cases where averages can be employed to greater advantage than any other form of figures. If, however, the Tables are to be used for further tabulation or inquiry rather than for comparison, the actual numbers should always be shown, even if they are approximated to a degree which does not affect the inferences to be drawn therefrom, in order that any further inquiry considered necessary or advisable can be carried out easily and correctly. When figures are large, and comparison consequently difficult, extra columns should be added, so that the percentages, or ratios which the items bear to the total, or the averages prepared on the lines laid down, may be shown. This is illustrated in the following Table :—

TABLE L.*

FREIGHT TRAIN TRAFFIC ON BRITISH RAILWAYS FOR THE MONTH OF JUNE.

Company.	General Merchandise.		Live Stock.		Fuel.		Other Minerals.		Total.
	Units.	% of Total	Units.	% of Total	Units.	% of Total.	Units.	% of Total	
G.W.R.	992.667	46	43,231	2	391,477	18	719,088	34	2,146,463
L. & N. E. R. ..	1,855,111	42	79,895	2	770,180	18	1,541,015	36	4,246,201
L. M. & S. R. ..	2,454,720	45	95,886	2	1,116,713	20	1,798,811	33	5,466,130
S. R.	539,741	53	11,245	1	110,574	11	351,592	35	1,013,152

* Compiled from the Ministry of Transport's Monthly "Railway Statistics."

When it becomes necessary to use actual figures, and these figures not only contain numerous digits, but are also large in number and tabulated in various columns, the difficulty of tracing the items on the same lines in such a compact mass of figures becomes very great, and the consequent confusion detracts considerably from the value of the table, and, moreover, the resultant tabulation does not appear attractive even to those whose business it is to study such returns. The Table shown on p. 124 is a summary of the Bank of England Return for fourteen weeks, in the form then presented by the *Economist.*

This Table is fairly easy to read, as the number of columns and the items are not very numerous, but as the number of columns and/or the number of items increases, the difficulty of reading the Table becomes greater and its value therefore less. Much, however, can be done to improve such masses of figures by care in the printing of them—by slightly widening the columns, by heavier division lines, and by leaving a blank space after each few lines (say after those figures relating to each month, or after every fifth line). The Table would then appear as on p. 125, and it will at once be seen that it is not only more readable, but its appearance is much more attractive.

The compiler must, however, study his problem carefully, and plan his table to meet the particular needs of each situation. A draft of the form should always be prepared before proceeding with the work of inserting the figures.

§ 5.—Title of the Table.

Having decided the use to which the Table is to be put, and the figures to be used to obtain the desired

TABLE M.

ANALYSIS OF THE BANK OF ENGLAND WEEKLY RETURN FOR THE FOURTEEN WEEKS ENDED 26TH JANUARY, 1927.

Date		Coin and Bullion.	Circulation (excluding Bank Post Bills).	Deposits.	Securities in Banking Department	Gold in or out of Bank.	Reserve and % of Reserve to Liabilities.		Bank Rate.
1926.		£	£	£	£	£	£	%	%
Oct.	27	152,814,940	139,069,400	122,606,011	106,809,136	1,260,000 out	33,495,570	$27\frac{3}{5}$	5
Nov.	3	152,807,082	139,537,365	122,226,842	106,901,562	28,000 out	33,019,717	27	,,
	10	152,060,534	138,834,485	120,801,594	105,576,781	727,000 out	32,976,049	$27\frac{5}{16}$	,,
	17	152,999,634	138,109,000	120,700,173	103,830,340	961,000 in	34,640,634	$28\frac{11}{16}$	,,
	24	152,974,616	138,004,975	124,634,409	107,698,730	11,000 out	34,719,641	$27\frac{7}{8}$	,,
Dec.	1	152,876,079	139,693,905	127,018,634	111,930,465	31,000 out	32,932,174	$25\frac{15}{16}$	,,
	8	153,233,705	139,634,485	120,390,704	104,877,660	431,000 in	33,349,220	$27\frac{3}{4}$	,,
	15	152,092,987	139,888,670	119,243,920	105,190,877	1,077,000 out	31,954,317	$26\frac{13}{16}$	,,
	22	151.943,215	141,285,355	120,074,136	107,616,958	158,000 out	30,407,860	$25\frac{5}{16}$	,,
	29	151,118,648	140,784,940	142,974,783	130,826,382	820,000 out	30,083,708	21	,,
1927.									,,
Jan.	5	151,380,637	139,803,280	152,583,667	139,300,786	228,000 in	32,327,357	$20\frac{1}{2}$	,,
	12	151,488,719	138,083,730	126.911,451	111,823,878	112,000 in	33,154,989	$26\frac{1}{8}$	,,
	19	151,491,156	137,100,715	123,801,072	107,759,555	28,000 out	34,140,441	$27\frac{9}{16}$	,,
	26	151,344,543	137,049,120	116,510,936	100,570,206	158,000 out	34,045,423	$29\frac{1}{4}$	,,

TABLE N.

ANALYSIS OF THE BANK OF ENGLAND WEEKLY RETURN FOR THE FOURTEEN WEEKS ENDED 26TH JANUARY, 1927.

Date.	Coin and Bullion.	Circulation (excluding Bank Post Bills).	Deposits.	Securities in Banking Department.	Gold in or out of Bank.	Reserve and % of Reserve to Liabilities.		Bank Rate.
	£	£	£	£	£	£	%	%
1926. Oct. 27 ..	152,814,940	139,069,400	122,606,011	106,809,136	1,260,000 out	33,495,570	$27\frac{3}{8}$	5
Nov 3 .	152,807,082	139,537,365	122,226,842	106,901,562	28,000 out	33,019,717	27	,,
10 ..	152,060,534	138,834,485	120,801,594	105,576,781	727,000 out	32,976,049	$27\frac{5}{16}$	,,
17 ..	152,999,634	138,109,000	120,700,173	103,830,340	961,000 in	34,640,634	$28\frac{11}{16}$	,,
24 ..	152,974,616	138,004,975	124,634,409	107,698,730	11,000 out	34,719,641	$27\frac{7}{8}$	,,
Dec. 1 ..	152,876,079	139,693,905	127,018,634	111,930,465	31,000 out	32,932,174	$25\frac{15}{16}$	,,
8 ..	153,233,705	139,634,485	120,390,704	104,877,660	431,000 in	33,349,220	$27\frac{3}{4}$	,,
15	152,092,987	139,888,670	119,243,920	105,190,877	1,077,000 out	31,954.317	$26\frac{13}{16}$	,,
22 ..	151,943,215	141,285,355	120,074,136	107,616,958	158,000 out	30,407,860	$25\frac{5}{16}$	,,
29 ..	151,118,648	140,784,940	142,974,783	130,826,382	820,000 out	30,083,708	21	,,
1927. Jan. 5 ..	151,380,637	139,803,280	152,583,667	139,300,786	228,000 in	32,327,357	$20\frac{1}{2}$	,,
12 ..	151,488,719	138,083,730	126,911,451	111,823,878	112,000 in	33,154,989	$26\frac{1}{8}$	,,
19 ..	151,491,156	137,100,715	123,801,072	107,759,555	28,000 out	34,140,441	$27\frac{9}{16}$	,,
26 ..	151,344,543	137,049,120	116,510,936	100,570.206	158,000 out	34,045,423	$29\frac{1}{4}$	,,

result, the next point of importance will be the **Title** of the Table. Since the principal idea in preparing Tabular Statements is to place data before people who have taken no part in their preparation, it is above all things necessary that the Table be entirely *self explanatory*, and for this reason great care must be exercised in the choice of the Title. This must be brief, but not so brief that its meaning becomes obscure or ambiguous. Due attention must also be paid to the Headings of the sub-divisions and columns, and when these are clear and well defined no direct reference need be made to them in the main title. The Title should therefore be wide in scope as well as clear in meaning. Wherever possible, the sub-divisions should be arranged in vertical form, and follow some well defined and easily understood order, *e.g.*, alphabetical, geographical, size or merit; the horizontal columns being utilised for units of measurement, when these are present. This form cannot, of course, be always followed, but experience will always indicate the form of Table most suited to the needs of the particular business.

In the general arrangement of the Table much can be done to assist the reader by the careful selection of type matter. The main heading should preferably be in the Roman type, and the size of the type used should vary with the importance of the division, and sub-divisions, all headings of equal importance being printed in the same style and size of type.

The Tabular Statement on page 127 pays due attention to all these points, and may therefore be taken as a model.

Generally speaking, a Typewriter is not suitable for tabulated statements, owing to the fact that the style and size of the characters is fixed to one type; but

TABLE O.*

TOTAL TON-MILES OF THE BRITISH RAILWAY COMPANIES FOR THE MONTH OF JUNE.

Railway Company.	Ton-Miles. General Merchandise (including Live Stock). No.	Per cent	Coal, Coke and Patent Fuel. No	Per cent.	Other Minerals. No.	Per cent.	Total.
Great Western	**77,002,269**	**58·87**	**19,891,854**	**15·21**	**33,907,369**	**25·92**	**130,801,492**
Western Area	33,765,617	64·50	5,625,373	10·75	12,956,134	24·75	52,347,124
Midland Area	25,125,927	53·39	6,487,727	13 79	15,448,182	32 82	47,061,836
South Wales Area	18,110,725	57·69	7,778,754	24·78	5,503,053	17·53	31,392,532
London & North Eastern..	**132,100,874**	**57·69**	**30,202,945**	**13·19**	**66,686,572**	**29·12**	**228,990,391**
Southern Industrial Area ..	43,415,161	58·24	7,748,765	10 39	23,407,476	31·37	74,607,402
Southern Agricultural Area..	25,094,366	56·77	2,690,597	6 09	16,414,267	37·14	44,199,230
North Eastern Area.. ..	39,703,664	59·86	8,342,505	12·58	18,281,701	27 56	66,327,870
Southern Scottish Area ..	22,343,212	53·51	11,102,826	26·59	8,311,184	19 90	41,757,222
Northern Scottish Area ..	1,508,471	71·88	318,252	15 16	271,944	12 96	2,098,667
London Midland & Scottish	**179,599,441**	**58·25**	**36,599,642**	**11·87**	**92,119,788**	**29·88**	**308,318,871**
Western Area, South ..	46,700,380	54·95	8,519,014	10 03	29,764,054	35·02	84,983,448
Western Area, North ..	40,551,421	56·60	10,520,724	14 68	20,578,952	28 72	71,651,097
Midland Area	62,902,560	59·38	11,039,189	10·42	31,997,688	30·20	105,939,437
Northern Area, South ..	22,037,460	62·56	4,894,959	13·89	8,296,184	23·55	35,228,603
Northern Area, North ..	7,407,620	70 44	1,625,756	15·46	1,482,910	14·10	10,516,286
Southern	**28,166,844**	**61·45**	**4,861,339**	**10·61**	**12,808,650**	**27·94**	**45,836,833**
Cheshire Lines Committee ..	5,563,111	64·45	792,670	9·18	2,276,689	26 37	8,632,470
Metropolitan	218,754	57·64	15,195	4·00	145,593	38·36	379,542
Midland & G.N. Joint ..	1,363,235	54·94	63,902	2·58	1,054,002	42·48	2,481,139
Great Britain .. .	428,212,813	58·38	93,204,993	12·71	212,058,795	28·91	733,476,601

* Compiled from the Ministry of Transport s Monthly "Railway Statistics."

where the Tables follow a pre-arranged form, and the form is a printed one, the Typewriter may be used with advantage in filling in the printed form—particularly as by means of a Tabulating attachment the figures can be correctly ranged, *i.e.*, units under units, tens under tens and so on. An adding attachment will also save much time in checking Totals. There is, however, one typewriter which is extremely suitable for Statistical work. This is the *Hammond Variotyper*, which has instantly changeable type faces, so that varied kinds of type can be used on the same form. It also embodies an arrangement whereby the spacing between the letters can be varied, in order that ten, fifteen or eighteen letters to the inch may be written. These two factors permit its use for even the most intricate Tabular work, as the writer has proved in actual practice.

§ 6.—The Uses of Percentages and Averages in Tabular Work.

In commercial statistics percentages will play a very large part, for since they are obtained from the condensation of large masses of data, seldom consist of more than three figures, and more generally two, it follows they are more easily read, and their significance and importance are grasped more quickly than would the figures upon which they are based. They enable comparison of various factors to be carried out easily, provided always that such factors are capable of comparison. It will usually be found advantageous to incorporate the percentages in the table with the actual figures, and thus ensure that wrong conclusions cannot be drawn.

The following table showing the daily sales of four branches of a commercial house for a week, enables two percentage calculations to be made, viz., the percentage of sales for each Branch, and the percentage of daily sales, both of course based upon the total sales of the whole period.

BRANCH SALES FOR THE WEEK ENDED 10TH FEBRUARY.

DAY.	Oxford	Bristol.	Norwich.	Derby.	TOTAL.	Percentage of Daily Sales to Total.
	£	£	£	£	£	
Mon., 5th Feb. ..	500	200	350	150	1,200	17·52
Tues., 6th Feb. ..	400	250	150	175	975	14·23
Wed., 7th Feb. ..	350	275	400	125	1,150	16·79
Thurs., 8th Feb...	550	225	425	130	1,330	19·41
Fri., 9th Feb. ..	200	250	375	170	995	14·53
Sat., 10th Feb. .	500	300	150	250	1,200	17·52
Branch Totals ..	£2,500	£1,500	£1,850	£1,000	£6,850	100
Percentage of Branch Sales to Total ..	36·49	21·89	27·01	14·61	100	

Such a statement is useful for comparing the ratio of daily sales, and also the total Branch sales. It can be extended over any period it is desired, and steady diminution in the percentage shown by any Branch would show that it was either falling off, or not sharing in the general progress. Thus investigations would be set on foot with a view to discovering the reason for the state of affairs revealed. Similarly, steps could be taken with a view to increasing sales on those days when sales were regularly lowest, such as special offers on such days, extended advertising, etc. While not showing all the details in percentage form, yet the percentages themselves supply a better guide to the real position than do the actual figures, for the ratios are not easily seen except by a trained mathematical observer.

Whenever calculations are necessary, great care should be exercised in checking the correctness of each stage of the work, preferably by some person other than the one who made the original calculation, for if errors be discovered by the reader of the Table doubt is thrown upon the whole of the Table, and its utility is destroyed or greatly impaired. It is particularly necessary to see that percentages are correctly worked, and averages prepared upon the correct bases. Alternative checks should be used whenever possible. When working percentages of various items of the same base the calculations can be checked by adding the percentages, since they should collectively amount to 100 per cent.; for instance:—

Branch.	Amount of Turnover.	Percentage of Total.
	£	Per Cent.
A . ..	60,000	20
B	75,000	25
C .. .	120,000	40
D	45,000	15
Total ..	**300,000**	**100**

It must, however, be remembered that when the percentages are worked out to fractions, and decimals are used correct to a defined place, it may happen that the result will not exactly equal 100 per cent., though it will always be within a very small fraction thereof.

If, however, the bases upon which the percentages are calculated vary with the different items, no such result can be obtained, so care must be taken to see that an error is not made by "taking the

average of an average." This is illustrated in the following example :—

Branch.	Turnover.	Expenses.	Percentage of Expenses to Turnover.
	£	£	Per Cent.
A	60,000	36,000	60
B	75,000	60,000	80
C	120,000	90,000	75
D	45,000	30,000	66·66
Total ..	**300,000**	**216,000**	**72**

To obtain the percentage which the total expenses of the business bear to the total Turnover it is necessary to work with the totals, *i.e.*, $\frac{216,000}{300,000} \times 100 = 72$ per cent., whereas a careless non-mathematical worker might add the percentages together and divide by the number of items, *i.e.*, $\frac{60 + 80 + 75 + 66\frac{2}{3}}{4} = 70{\cdot}416$ per cent., a result which is entirely wrong, and which if used for subsequent work would lead to further serious errors.

Commercial men frequently use statistics for the purpose of maintaining or increasing sales, and in this connection probably the percentage statement combined with the actual figures is the simplest form in which to present to the Sales Staff the result of their labours. In many fields of salesmanship at the present time individual salesmen or branches are allocated a "quota" of business to be done. Such quota is fixed more or less scientifically, following an investigation of the nature of the area covered, the possible purchasers, the business previously done, and the ability of the representative. The following table

prepared monthly will show exactly what is being done, and the salesman can see whether he is selling the proportion of the goods necessary for him to reach his quota at the end of the year :—

SALESMAN'S (OR BRANCH) STATEMENT OF BUSINESS
FOR MONTH ENDED 31ST JANUARY.

Area or Salesman.	Quota for Year	Business done to date.	Percentage of Quota.
	£	£	
A	10,000	1,000	10
B .	8,000	500	6·25
C	12.000	1,000	8·3
D . .	12,000	800	6 6
E .	18,000	1,700	9·$\dot{4}$
	£60,000	£5,000	8·$\dot{3}$

If the business were one where trade was steady throughout the year, then 8·$\dot{3}$% of the quota is the average amount per month required to ensure success. The above statement shows that the monthly average was obtained, but that Salesmen B and D are below this standard and hence would need to work harder. A and E on the other hand are well above their average, a satisfactory commencement to the period under review.

At the end of the trading period a statement on the following lines would show the position at a glance :—

SALESMAN'S (OR BRANCH) STATEMENT OF BUSINESS DONE
FOR YEAR ENDED 31ST DECEMBER.

Area or Salesman.	Quota for Period.	Actual Business Done.	Amount Above or Below* Quota.	Percentage Above or Below* Quota.
	£	£	£	
A	10,000	12,000	2,000	20
B	8,000	7,750	250*	3·125 *
C	12,000	13,000	1,000	8·$\dot{3}$
D	12,000	11,500	500*	4·16*
E	18,000	24,000	6,000	33·33
	£60,000	£68,250	£8,250	13 75

There are many variations of such statements in common use. Their advantages are many and their use being constantly extended.

Percentages incorporated into statements of account also enable comparison to be readily made, but as a general rule it is difficult to incorporate the necessary comparative percentages figures and yet maintain the simplicity of the table. The statement on page 132 shows an effort to combine ease of comparison with clearness and simplicity, and yet give the maximum of information to the reader.

It will be noticed that percentages are computed on two different bases, and this procedure is undoubtedly essential in Commercial Accounting work. The expenses incidental to the actual production of the goods, including the wages of the workers, and factory overhead expenses or oncost, should always be calculated on the basis of the cost of manufacture. All matters appertaining to the sales, and to the general administration of the business will be based on Turnover. This results in the Net Sales being equated to 100, and all the percentages of expenses are calculated on this base. If cost were taken, then the sales would be in excess of 100, and the expenses of salesmen and others on the selling and administrative staffs would be calculated not upon the income of the business, but upon a part only of the outgoings. The standard percentage might be based on the average or moving average of the figures for a number of years, or it might be the results of the previous year or some other base chosen by the compilers.

A series of such statements can be condensed into the form on page 135, which gives important comparisons.

A. B. & Co.

PROFIT AND LOSS STATEMENT FOR YEAR ENDED 31ST DECEMBER, 1935.

Percentage of Cost. Standard.	Percentage of Cost. Actual 1934.		£	£	£	Percentage of Turnover. Standard	Percentage of Turnover. Actual 1934.
		SALES			500,000		100%
		Deduct:					
		COST OF GOODS SOLD.					
		Stock of Raw Material 1st January, 1935	50,000				
		Purchases, 1935	270,000				
		Total	320,000				
		Stock of Raw Material 31st December, 1935.. ..	60,000				
60	63·41	Cost of Raw Material used:		260,000			
35	31·71	Factory Wages		130,000			
4	4·27	Manufacturing Expenses		17,500			
1	·61	Carriage on Raw Material		2,500			
100	100	COST OF GOODS MANUFACTURED		410,000			
		Deduct:					
		Difference in Stock of Finished Goods:					
		31st December, 1935 ..	35,000				
		1st January, 1935 ..	25,000	10,000			
		COST OF GOODS SOLD.			400,000	76	80
		GROSS PROFIT ON SALES ..			100,000	24	20
		Deduct:					
		Selling Expenses			10,000	4	2
		ACTUAL PROFIT ON SALES ..			90,000	20	18
		Deduct:					
		Administrative Expenses ..			7,500	2	1·5
		NET PROFIT ON TRADING..			£82,500	18	16·5

A. B. & Co.

Analysis of Condensed Profit and Loss Statements for Five Successive Years.

Year.	Net Sales	Cost of Goods Sold.	% of Cost to Sales	Gross Profit on Sales.	% of Gross Profit.	Selling Expenses.	% of Selling Expenses.	Actual Profit on Sales.	% of Net Profit on Sales.	Administrative Expenses.	% of Sales.	Net Profit on Trading.	% of Net Profit.
	£	£		£		£		£		£		£	
1	300,000	200,000	$66\ \dot{6}$	100,000	$33\cdot\dot{3}$	20,000	$6\cdot\dot{6}$	80,000	$26\cdot\dot{6}$	7,000	$2\cdot\dot{3}$	73,000	$24\cdot\dot{3}$
2	450,000	250,000	$55\ \dot{5}$	200,000	$44\cdot\dot{4}$	40,000	$8\cdot\dot{8}$	160,000	$35\cdot\dot{5}$	9,000	2	151,000	$33\cdot\dot{5}$
3	600,000	450,000	75	150,000	25	30,000	5	120,000	20	10,000	$1\cdot\dot{6}$	110,000	$18\cdot\dot{3}$
4	400,000	310,000	77 5	90,000	22 5	20,000	5	70,000	17 5	10,000	2·5	60,000	15
5	500,000	400,000	80	100,000	20	10,000	2	90,000	18	7,500	1·5	82,500	16·5
Average	£450,000	£322,000	$71\cdot\dot{5}$	£128,000	$28\cdot\dot{4}$	£24,000	$5\cdot\dot{3}$	£104,000	$23\ \dot{1}$	£8,700	$1\cdot9\dot{3}$	£95,300	$21\cdot1\dot{7}$

It will be seen that in the second year the Turnover increased without the Cost of the Goods sold increasing in like degree. Net profits showed a considerable increase although expenses also showed sharp advances. In the three last years the cost of production shows steady increases and a corresponding reduction in profits, as the price realised on the sales has not advanced in the same way as the cost. This may be due to competition preventing higher prices being charged for the goods, but the actual reason needs to be enquired into. In the last year expenses were materially reduced and turnover considerably increased, thus leading to a higher rate and larger actual profits, although the cost of production still showed an advance. By averaging the monetary columns, and working out the percentages on such average, a standard based on the five years working could be obtained as indicated in the Statement.

§ 7.—Tabulation of Commercial Data.

The use of Statistical matter need not necessarily be confined to the inside work of the office, but may frequently be used to support arguments advanced to customers or prospective customers. On the opposite page is an advertisement taken from an American paper, which illustrates the uses to which statistical records may be put.

From the point of view of the Statistician the statement, though clear and lucid, has several weak points. An instance is the inclusion of the number of days operated by the two trucks without assigning reasons for the wide divergence which appears in the number of days worked, as it is hardly possible that the time the Truck numbered 44 was not working

was directly attributable to tyre trouble, though if such happened to be the case the statement would be in order. The comparison of the miles travelled is also fallacious, in so far that the number of days worked by the Trucks was unequal, that is unless tyre trouble, or other troubles directly due to the tyres are responsible for Truck numbered 44 being off the road. If

ADVERTISEMENT.

Recently, a six months' test of solid versus pneumatic truck tyres was conducted by Mandel Brothers, operating a large downtown dry goods store in Chicago. The test was supervised by the company's chief engineer, Mr. Alfred Johnson. Two trucks were used, No. 44 on solid tyres, and No. 45 with Goodyear Cord Tyres on the rear wheels and other pneumatics on the front wheels. Both trucks were employed in regular city delivery work, handling similar loads over similar routes. The results of the test, given below, have caused the company to order recently several new trucks on pneumatics.

	Truck No. 44 (Completely equipped with Solid Tyres).	Truck No. 45 (Completely equipped with Pneumatic Tyres).	Difference in Favour of Pneumatic Equipment.
Number of days operated ..	130	146	12·3% increase
Miles travelled	5,388	6,719	25 % increase
Miles per gallon of gasoline ..	5·5	7·2	31 % increase
Repairs per mile	$ ·0109	$ ·0057	47 % saving
Labour Cost per mile (Drivers)	$ ·20	$ ·176	12 % saving
Operating Cost per mile ..	$ ·353	$ ·335	5 % saving

NOTE.—The actual saving in operating cost amounted to 1·8 cents per mile. Such a reduction grows to a very substantial sum of money when multiplied by thousands of miles of service.

we take the average mileage covered per day, we find that Truck No. 44 did $\frac{5,388}{130}$, or 41·45 miles per day, whereas No. 45 averaged $\frac{6,719}{146}$, or 46 miles per day. This gives an increase in favour of the latter of 11·097 per cent., a very different result from the 25 per cent.

increase shown in the Table. If, however, the Statement showed the reason for the difference in the number of days worked, criticism would not be so likely to arise.

This illustration will serve to emphasise the necessity for careful preparation of Statistical matter, the elimination of all matters of doubt, and the substitution of data or headings which will prevent criticism, and avoid confusion or doubt.

SYNOPSIS TO CHAPTER VI.

THE GRAPHIC METHOD.

§ 1.—Applications of the Graphic Method.

2.—The Frequency Polygon.

3.—The Normal Frequency Curve.
- (a) Skewness.
- (b) Smoothing the Curve.
- (c) When Data consists of items of known magnitude.

4.—The Ogive.

5.—Kinds of Diagrams.
- (a) The "Block" Diagram.
- (b) The "Bar" Diagram.
- (c) The Graph.
  - (1) The Natural Scale Graph.
  - (2) The Logarithmic Scale Graph.
- (d) The "Bar" Diagram to show Components.
- (e) Comparisons of Areas and Volumes.
- (f) Diagrams to show three Factors.
- (g) Diagrams showing two Scales.

6.—Advantages and Disadvantages of the Graphic Method.

7.—Rules for Constructing Diagrams.

CHAPTER VI.

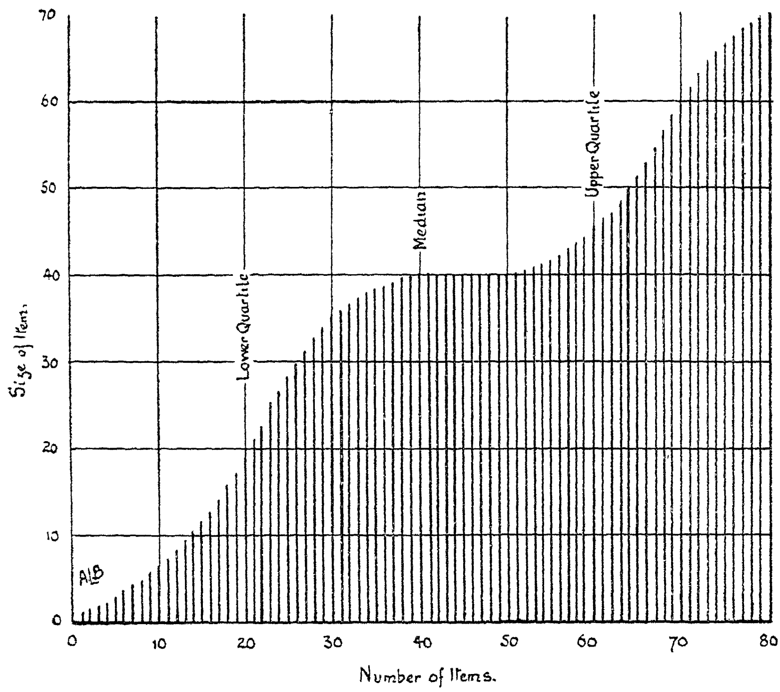
THE GRAPHIC METHOD.

To the trained mind a carefully compiled Table is of great value, both for purposes of reference and for deductive reasoning, but if Statistical methods are to become of really great utility to the business man who is not a skilled mathematician, it behoves us to present our data and facts in such a manner that not only are the principal details supplied, but the entire meaning and trend of the data presented can be seen without the necessity for long consideration and study. It must be frankly admitted that Tabulation falls very short of our Standard in this respect, for though the meaning can be grasped by one with mathematical training, it only presents a mass of figures to others. Tabulated statements, however carefully prepared, are not attractive to the business man, who finds the study of them tedious and involved, so that frequently it is necessary to call for explanations and reasonings from those responsible for their compilation. It is difficult for those not accustomed to mathematics to visualise conditions and results from mathematical formulæ and data. The actual trend of the figures is frequently obscured when the data is presented in Tabular form, and the eye is frequently tired by referring to masses of figures. PICTORIAL REPRESENTATION, when properly carried out, appeals to the eye,

and also to the mind, because it is practical, clear, and easily understandable even by those unacquainted with the method of preparation. A statement that a series of articles are arrayed in order of magnitude with the smallest on the left and the largest on the right does not convey the same result to the mind as is obtained by glancing at the same facts presented in Diagram No. 1. While an effort has to be made to visualise the position when only the written statement is available, a mere glance at the diagram is sufficient to make the meaning clear, and time is thus saved to the busy man, while the possibility of errors arising from wrong impressions is also reduced considerably. A busy physician visiting patients in hospitals obtains a very easily-read clue to the progress and condition of his patient by a glance at the chart of his temperature which usually hangs at the head of each bed. This chart tells him the changes which have taken place since his last visit, and eliminates the necessity for inquiry as to what has happened in his absence. The wandering of a line is more powerful in its effect on the mind than a tabulated statement; it shows what is happening, and what is likely to take place, just as quickly as the eye is capable of working. The graphic method so far has not been used very largely in commerce, but its utility is being increasingly recognised, for its advantages are many. Few people, however, are accustomed to read graphs, and thus do not appreciate their usefulness. As with Tabulation, great care must be taken in the preparation of diagrams and graphs, though they are capable of being adapted to any circumstances and needs which may arise. *Diagram No.* 1 shows the simplest possible form of diagram. It is easy of compilation, requires no particular skill or training to

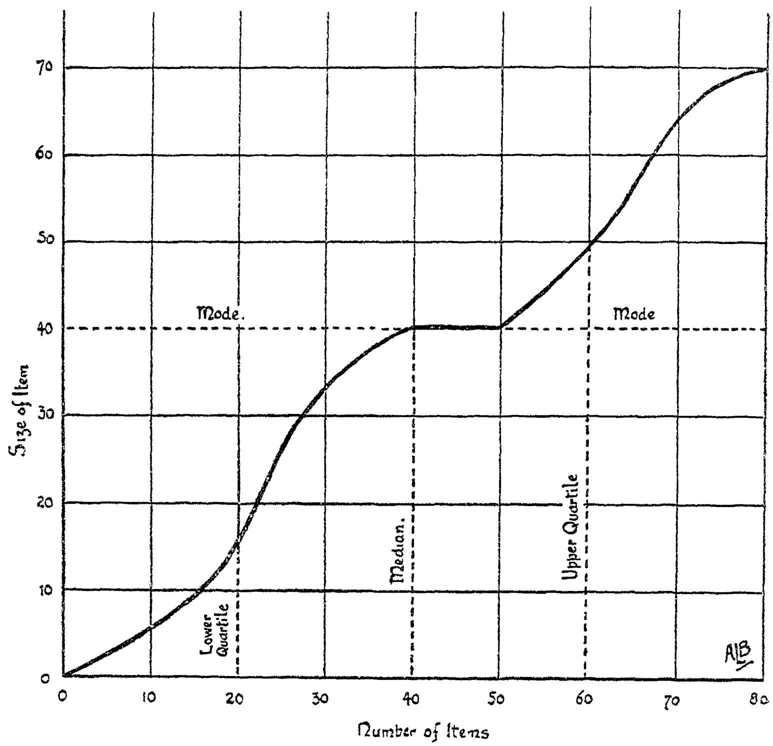
DIAGRAM No 1.

SHOWING A SERIES OF EIGHTY ITEMS ARRAYED.



construct, is easily read and understood, and the significance of the various factors is at once apparent. It is not, however, an ideal method of diagrammatic presentation, for it involves considerable work, and great care is needed in drawing the many lines in the data; while if these lines become very numerous it would be necessary either to enlarge the diagram, or else draw the lines closer together. Both of these alternatives would result in the eyes becoming confused with the multitude of lines, the diagram consequently declining in value. Such data are always better plotted in the form of a curve, as is shown in *Diagram No.* 2. The magnitudes of the items are marked on the sectional paper in the same manner as in the first example, but, instead of drawing lines from the base line to the points so obtained, these latter are joined together by a curve, which gives exactly the same outline as Diagram No. 1. Owing to the absence of the numerous lines the varying sizes of the items can be traced more easily than in the first diagram, and the representation is much clearer. Such a curve is easily plotted and drawn, and, moreover, if the items are more numerous than those shown, the horizontal scale can be adjusted and the curve re-drawn without the chart becoming more confusing or unreadable, and this is a very distinct advantage. Though this particular type of diagram or Histogram (*i.e.*, the plotting of an array of the data) is not likely to be met with to any degree in the commercial world, the principle upon which it is prepared is an extremely useful one, and capable of almost infinite variations, being thus adaptable to any particular data or problem with which one may be confronted.

DIAGRAM No 2
Curve showing "Array" of Eighty Items of varying Size.



§ 1.—Applications of the Graphic Method.

Before considering the utility of the graph to the business man it will be well to consider how it is applied to the science of Statistics, and the lessons thus learnt can be applied to the everyday problems of the commercial world. We have already seen that if we collect data of any description and grade them, there will be found a tendency for the items to fluctuate around a particular type which we already know as the Mode. This is true of all chance and natural phenomena. The majority of the items will be found to be grouped near this Mode, and as the distance from the Mode increases in either direction the number of items becomes less, and, in the case of natural phenomena, approximate limits within which fluctuations occur can be fixed. As has already been seen in Diagram No. 1, and as is shown even more clearly in Diagram No. 2, the Mode is located in the series of data plotted therein, between the fortieth and forty-ninth items, since they are all of the same magnitude, and the curve "flattens" out at this point in the array after rising sharply and continuing with equal steepness. Experience shows that it will make no difference to our result if we increase the number of items under review, *provided that those samples we have taken and plotted are truly representative of the whole of the examples.* It will thus be seen that the correctness of our work in the early stages of the compilation is of supreme importance, if our deductions are to be of real use.

Usually, however, it will be found that the data collected do not take the simple form shown in the diagrams numbered 1 and 2, but many of the items are of equal dimensions. When this is the case it becomes necessary to compile and plot a frequency table, and

it is then essential that the class interval (*i.e.*, the variations in the magnitude of the data) should be uniform. Let us assume that a census has been made of the businesses of a particular kind of trade, and from the results we have selected 335 examples where the capital involved is the same, but the profits earned vary from £1 to £4,500. Here the variation between the smallest and largest examples is very wide, and we get a great number of items between the two extremes varying so little in magnitude that to plot them all correctly would mean a very big and cumbersome diagram. In such cases the only thing to do is to take arbitrary dividing lines and collect the examples into groups, taking care that the class interval is the same in each case. Let us assume that the following simple Frequency Table is obtained as a result of such grouping :—

TABLE P.

TABLE SHOWING FREQUENCY WITH WHICH CERTAIN DATA OCCURS.

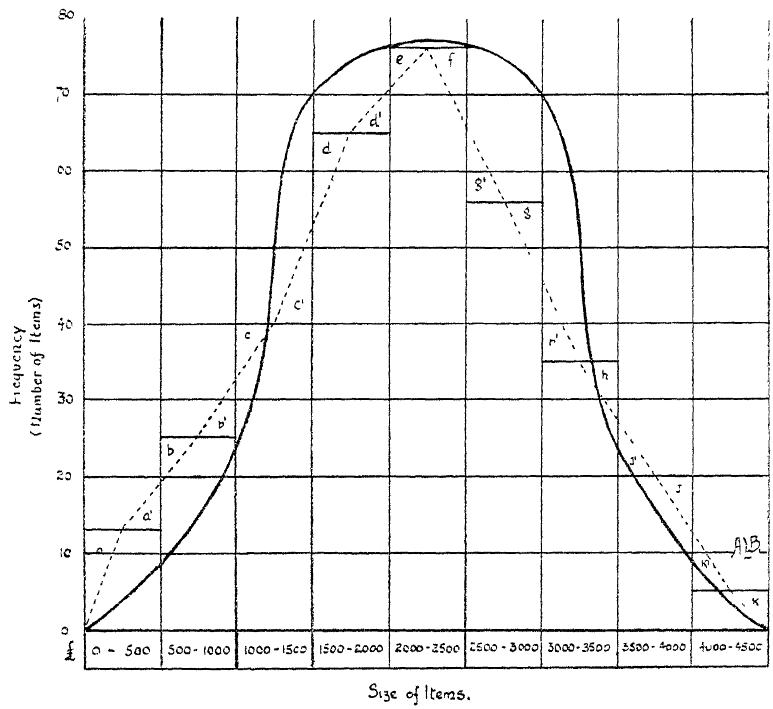
Size of Item. (Amount of Profits.)	Frequency. (Number of Examples in Each Group.)
Not exceeding £500	13
£501 to £1,000	25
£1,001 ,, £1,500	40
£1,501 ,, £2,000	65
£2,001 ,, £2,500	76
£2,501 ,, £3,000	56
£3,001 ,, £3,500	35
£3,501 ,, £4,000	20
£4,001 ,, £4,500	5

§ 2.—The Frequency Polygon.

This information would then be plotted as in Diagram No. 3, with rectangular blocks to represent the numbers in each group. As the class boundaries are chosen in an arbitrary fashion, it follows that the rectangular diagrams would be entirely different were the examples grouped with other boundaries instead

DIAGRAM No 3

Diagram showing Frequency when Items are grouped



of those shown. If the groups were, for instance, to be limited to intervals of £250, the rectangles would be narrower and the height of the steps would decrease, and if this process were continued sufficiently far a smooth curve would result, but such a result would cause nearly as much work as plotting the actual magnitude of all the examples. It must also be recognised that within each of the groups plotted a variety of magnitudes will be present, and therefore the rectangles do not present a true picture. We can, however, approximate the result which would be obtained by plotting each individual item by connecting the outer points of the base of the diagram (*i.e.*, 0 and £4,500) through the middle points at the *top* of each of the rectangles as shown by the dotted line on Diagram No. 3. The resultant figure is known as a "**Frequency Polygon,**" and it will be seen that the area of such Polygon is approximately the same as that of the rectangles. In the diagram under consideration the triangles outside the new figure, which are included in the original area, roughly correspond in area with those now included, but which were previously excluded (compare b and b^1, c and c^1, etc.), but it will be noticed that the triangles e and f are excluded entirely, and that the area of the Frequency Polygon is therefore a trifle smaller than that of the area enclosed by the rectangles. It must also be recognised that other items might possibly occur outside the limits chosen, *e.g.*, a case might occur of profits amounting to more than £4,500, and if some other class interval had been taken there is a possibility that the rectangle for the entire group might be higher, and as a consequence the apex of the polygon might be above that now shown. Moreover, if our data be truly representative then the collection of additional

examples should result in the frequency of the Modal size (*i.e.*, the group in which the profits ranged from £2,001 to £2,500) being increased to a greater extent than any other group, and this would also result in the apex of the polygon being above that shown in the diagram. Even though the data shown in the Frequency Polygon be representative, yet it must be borne in mind that if additional examples be collected subsequent to its construction a large proportion of such items will tend to fall into the Modal group, or into the groups immediately adjacent to it, and hence the extremes will decline in importance compared with such groups.

§ 3.—The Normal Frequency Curve.

Bearing these facts in mind, we can now draw a final smooth curve, and this will show what is believed to be the actual distribution of the profits earned in that trade with that particular amount of capital, assuming that all the examples in existence were actually included. This curve will be seen from the diagram to be bell-shaped, and is known as the **Normal Frequency Curve**, or occasionally as the **Normal Curve of Error.** In constructing such a curve we endeavour to eliminate all accidental variations which appear in the samples taken and plotted, and thus obtain a result which in all normal circumstances might be regarded as the correct distribution of the data. Such a bell-shaped curve can be constructed for *all* chance or natural phenomena, but in other data, such as those appertaining to economic or sociological investigations, the examples may display such wide irregularities that we may find it impossible to obtain such a Normal Curve of Error. This would to a very large extent be true also of commercial data, for though two businesses in the same trade may be commenced at the same time,

with the same capital, and with the same prospects of success, the results may be, and often are, widely divergent, such divergence being due to one cause, or a combination of many causes of a varying nature, such as a better business position, greater personal skill of the proprietor, better methods of management, more facilities for credit, etc.

(a) **Skewness.**

When the Frequency Table shows that the distribution of the data is not even, *i.e.*, that the number of examples falling into classes above and below, but equi-distant from the Modal Group are not approximately the same, then the Curve of Normal Error will be distorted. When this happens and the two sides of the bell-shaped curve are not symmetrical there is said to be **Skewness**. In Table P (page 146) the data is distributed fairly evenly, but had the frequency been as follows the curve would have appeared as in Diagram No. 3a.

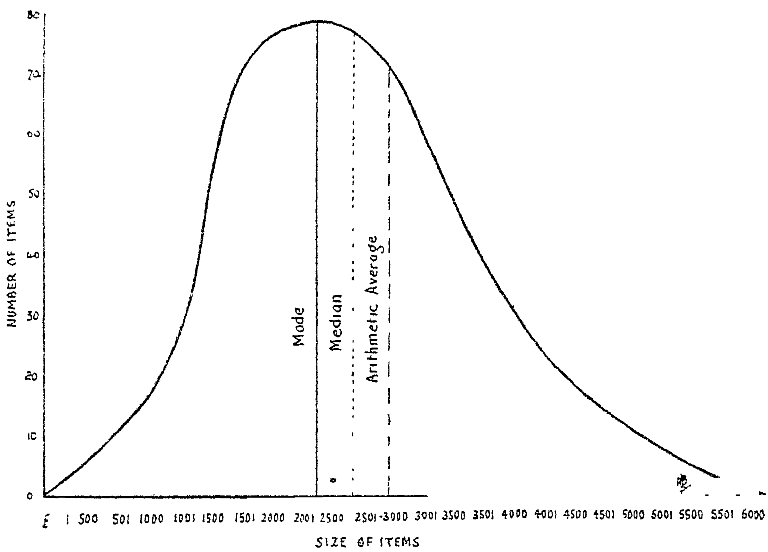
TABLE Q.

Size of Item. (Amount of Profits.)	**Frequency.** (Number of Examples in Each Group.
Not exceeding £500	13
Exceeding £500 and not exceeding £1,000 ..	25
,, 1,000 ,, ,, ,, 1,500 ..	40
,, 1,500 ,, ,, ,, 2,000 ..	65
,, 2,000 ,, ,, ,, 2,500 ..	76
,, 2,500 ,, ,, ,, 3,000 ..	65
,, 3,000 ,, ,, ,, 3,500 ..	56
,, 3,500 ,, ,, ,, 4,000 ..	45
,, 4,000 ,, ,, ,, 4,500 ..	32
,, 4,500 ,, ,, ,, 5,000 ..	16
,, 5,000 ,, ,, ,, 5,500 ..	10
,, 5,500 ,, ,, ,, 6,000 ..	6

The curve is materially distorted by the magnitude of the additional data which has come under review, the additional 114 examples taken all falling into groups larger than the Mode, without however, affecting its position.

CHART No. 3a

Curve of Normal Error showing Skewness



Whenever the data is regularly distributed, and the Curve of Normal Error takes the form of a bell-shaped curve, it will be found that the Median will fall within the Modal Group, for it always bisects the area of the diagram. This happens in the data shown in the Frequency Table P. The Median is the 168th item, *i.e.*, the 25th item in the Modal Group (size £2,001 to £2,500). In the data given in Table Q, however, the Median is the 225th item and this falls in the group of size £2,501 to £3,000. The result therefore of the skewness has been to shift the Median to the right, and to bring it outside the Modal Group, thus making the Median less representative than it would be were the distribution even. Similarly, the Arithmetic Average will have changed, as this is always located at the centre of gravity in such diagrams. If we assume that the data in each of the groups in the two tables P and Q are evenly distributed, a weighted descriptive average can be obtained which will demonstrate this fact.

			(a) Data in Table P.		(b) Data in Table Q.	
Group No.	Magnitude of Items.	Average Profits of Examples in Group	No. of Examples in Group	Product.	No of Examples in Group.	Product.
		£				
1	Not exceeding £500 ..	250	13	3,250	13	3,250
2	£501 to £1,000	750	25	18,750	25	18,750
3	£1,001 ,, £1,500 .. .	1,250	40	50,000	40	50,000
4	£1,501 ,, £2,000	1,750	65	113,750	65	113,750
5	£2,001 ,, £2,500	2,250	76	171,000	76	171,000
6	£2,501 ,, £3,000	2,750	56	154,000	65	178,750
7	£3,001 ,, £3,500	3,250	35	113,750	56	182,000
8	£3,501 ,, £4,000	3,750	20	75,000	45	168,750
9	£4,001 ,, £4,500	4,250	5	21,250	32	136,000
10	£4,501 ,, £5,000	4,750			16	76,000
11	£5,001 ,, £5,500	5,250			10	52,500
12	£5,501 ,, £6,000	5,750			6	34,500
			335	720,750	449	1,185,250

$$\text{Descriptive Average in } (a) = \frac{720{,}750}{335} = £2{,}151{\cdot}492$$

$$\text{Descriptive Average in } (b) = \frac{1{,}185{,}250}{449} = £2{,}639{\cdot}755$$

The average has therefore increased by no less than £488·263.

In each case, however, the Mode remains unchanged, but both the Median and the Arithmetic Average have moved materially upwards, and such changes would naturally result in different conclusions being drawn when either the Median or Arithmetic Average were being used as a base for comparison with other data, as against those drawn when the Mode was taken as the base.

It will be found that as a result of this skewness the Arithmetic Average has moved more than the Median, and such is always the case. The magnitude of the Median in Table P is:

$$£2{,}001 + \left(\frac{25}{76} \text{ of } £499\right) = £2{,}165{\cdot}14$$

While in Table Q it is:

$$£2{,}501 + \left(\frac{6}{65} \text{of } £499\right) = £2{,}547{\cdot}06$$

a movement of £381·92 only, a marked difference from the movement in the Arithmetic Average shown above. Skewness will always result in the Mode, the Median and the Arithmetic Average following one another in that sequence, the Arithmetic Average being moved the furthest on that side in which skewness has developed.

(b) Smoothing the Curve.

When plotting our data it is not always necessary to plot the rectangles first, as a point can be plotted which would be identical with the middle position at

the top of the rectangle, and thus the first thing to appear would be the Frequency Polygon. In drawing the smoothed curve care must be taken to see that the curve changes direction as little as possible, and all irregularities should be smoothed out. The extent to which the smoothing will be carried will depend upon the data involved, and no hard-and-fast rules can be laid down upon the matter; but since almost every smoothed diagram representing a continuous series will commence with an extremely small number of instances, and after reaching its maximum decrease slowly to zero, the curve should naturally begin and end on the base line.

(c) When Data consists of Items of Known Magnitude.

It may sometimes happen that the information at our disposal is of such a nature that the actual magnitudes of the examples can be plotted, and we shall find that if we pursue similar methods we shall obtain a very similar curve. Let us assume that the Frequency Table on the opposite page shows the *actual* profits earned in the 335 cases previously taken. From these data we can construct a Frequency Line or Bar Diagram as shown in Diagram No. 4. The size of the items (in this case the profits earned) are plotted on the horizontal scale, while the frequency with which they occur is placed on the vertical scale, and a line drawn from the base line to the point on the vertical scale corresponding to the number of cases of that particular magnitude which occur in the samples under review. These lines are those marked *a* in the diagram. A study of the diagram makes it apparent that the Modal profit is £2,250, for there are no less than forty cases in the Table under review. If the law which has been enunciated above applies, as it will to some extent at least, it would follow that if these are really

representative examples of the whole, then we may expect that all other examples would follow the same lines, but to guard against error in our assumptions, and bearing in mind that the Modal Group and those adjacent thereto will be emphasized if additional examples be collected, we usually smooth the diagram by drawing a curve to allow for any possible contingencies. When this has been done it will be seen that the curve is again of a bell shape. In the case illustrated in Diagram No. 3 we were aware that there were intermediate examples, but in this case there are no examples

TABLE R.

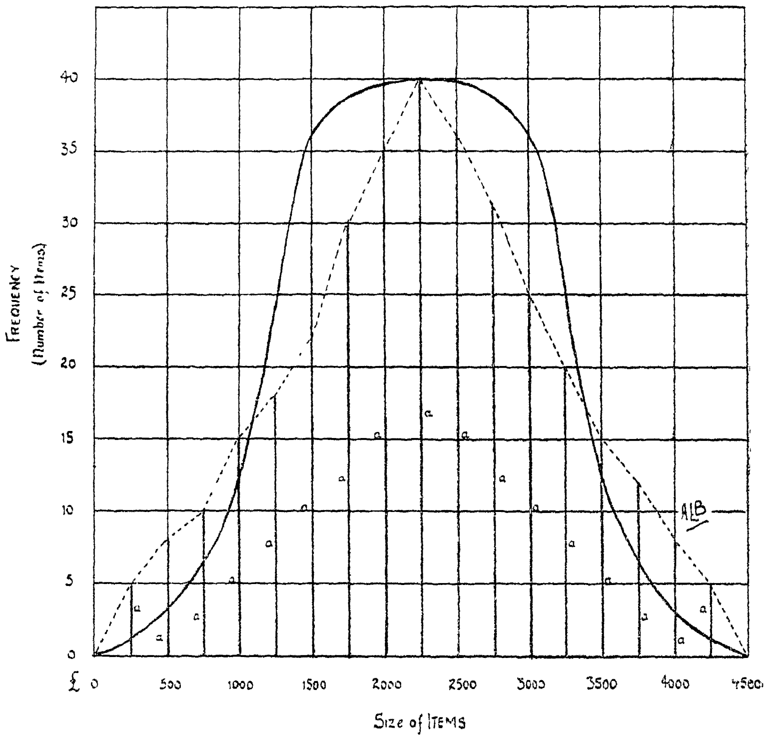
TABLE SHOWING FREQUENCY AND CUMULATIVE FREQUENCY OF DATA.

Size of Item. (Amount of Profits.)	Frequency. (Number of Examples.)	Cumulative Frequency.
£		
250	5	5
500	8	13
750	10	23
1,000	15	38
1,250	18	56
1,500	22	78
1,750	30	108
2,000	35	143
2,250	40	183
2,500	36	219
2,750	31	250
3,000	25	275
3,250	20	295
3,500	15	310
3,750	12	322
4,000	8	330
4,250	5	335
4,500	0	335

known which fall between those plotted, consequently in this case the curve must not be interpreted as showing the existence of intermediate examples as in the last case, *but only as a curve of tendency*, and in reality the curve therefore represents the relative height to which the vertical lines would extend were the number of examples infinite.

DIAGRAM No. 4.

Diagram showing Frequency with which Items occur

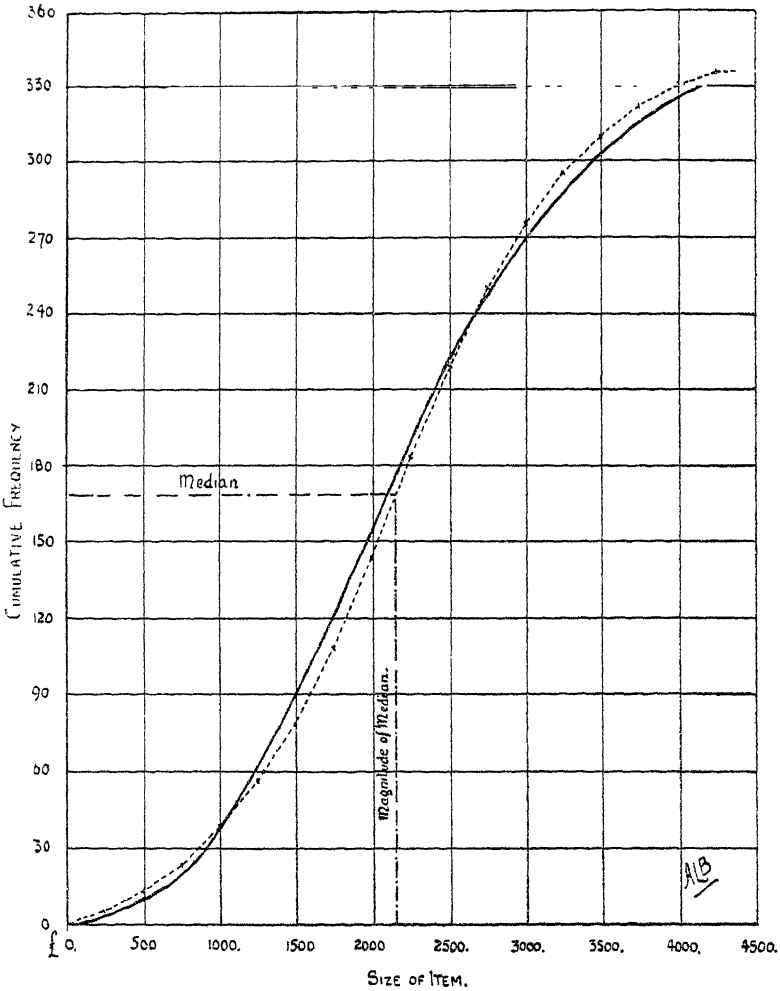


The advantages of such a curve of Normal Frequency are numerous, for by its aid we can visualise what is happening to particular sections of trade and commerce, and so estimate the possibilities of profit from such sections. If such curves were obtainable, and were carefully studied by those commencing business, the result would be that speculation would be to a large extent eliminated, and efforts made to obtain the Modal result shown by the graph, instead of remaining satisfied with a smaller return. This would tend to lead to more efficient management.

§ 4.—The Ogive.

Another type of diagram used for the consideration of data into which no time element enters is illustrated in Diagram No. 5, and the resultant curve is known as an "**Ogive.**" This is obtained by plotting the *cumulative frequency* of a series of observations, and in this particular example the cumulative frequency of the profits enumerated in Table R (*supra*) have been taken. The dotted line shows the original Ogive plotted from the actual figures, while the continuous thick line is the smoothed curve resulting from the plotting of the data in question. It is usually much easier to smooth an Ogive than a Frequency Polygon, and the Ogive is very convenient for locating the Median Quartiles, Deciles, etc. To find the magnitude of the Median it is necessary to locate the number on the Cumulative Frequency Scale (*i.e.*, the vertical scale) and draw a horizontal line from this point to the curve, and then from the point so obtained drop a vertical line to the horizontal base line (or axis), and the point where the vertical line intersects such base will give the magnitude of the Median example. By means of such graphs the size of the Median can be

DIAGRAM No 5
"OGIVE" to show Cumulative Frequency.



accurately obtained even when the data is indefinite in some respects, and this method is easier than the arithmetical calculation previously explained in Chapter IV. It is also possible by means of a similar operation to that described above to obtain the magnitude of any particular item under review, while if it is desired to obtain the number of examples not exceeding a certain size the information can be obtained by raising a perpendicular from the base point coinciding with the limit of magnitude specified. From the point where this perpendicular meets the Ogive a horizontal line is run to the left hand margin and the scale there gives the number of examples. The number of examples in any group or combination of successive groups can also be obtained by means of perpendiculars being raised at the Upper and Lower Magnitudes of the groups selected. The horizontals will then give two points on the scale, and the difference between the numbers so obtained will be the number of examples in the groups selected.

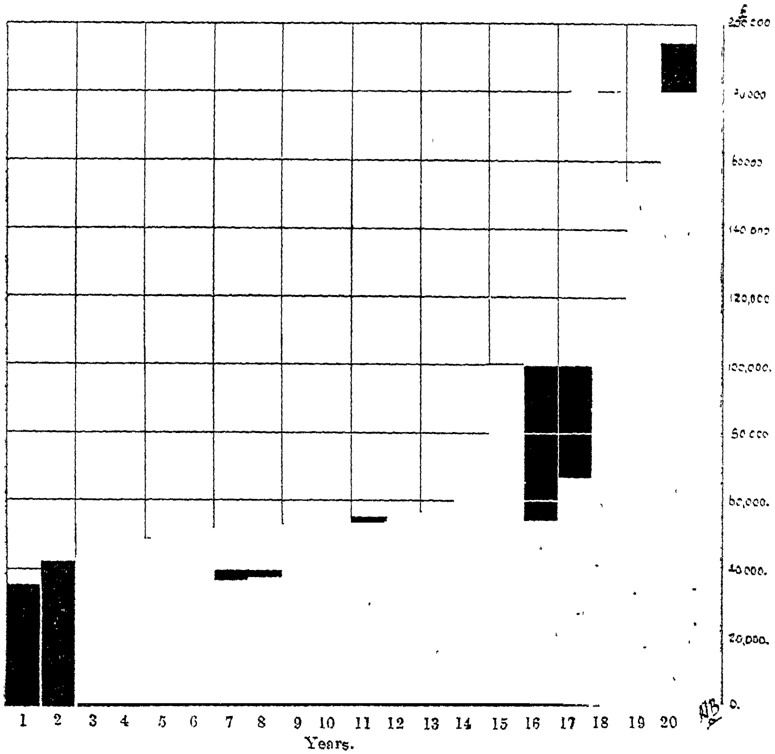
It will be observed that we cannot ascertain the Arithmetic Average of the series by inspection of either of the curves appearing in Diagrams 2, 3, 3a, 4 or 5, whereas the Median, Quartiles, Deciles, or other similar types, are easily located and their magnitude measured by a simple mechanical process which does not lend itself to error. The Mode cannot be located easily on the Ogive, but would be at that part of the curve which is steepest.

§ 5.—Kinds of Diagrams.

So far we have only considered that form of data into which no element of time enters, but in considering most of the forms of commercial data, particularly when comparison becomes necessary, time enters into

DIAGRAM No 6.

Diagram showing Annual Turnover for Twenty Years



the data and plays a conspicuous part, and hence our graphs must take a somewhat different form. Such Diagrams are sometimes known as "**Historigrams.**"

(a) The "Block" Diagram.

We have recently seen efforts made at presenting commercial data in graphic form, particularly in connection with company flotations, but so far no effort at uniformity in method has been made, and many of the diagrams are faultily conceived and badly executed. Diagram No. 6 is based upon the method used in one such published diagram, and may be termed the "Block" method of presentation. This diagram shows the Turnover of a business for twenty years from the figures shown on page 77, and though it is a step in the right direction, the diagram is clumsy both in conception and execution. It is difficult to read, for since the co-ordinate lines are obscured, the eye has great difficulty in following the heights of the various rectangles to the scale which appears on the right-hand side of the diagram. This scale, too, is wrongly placed, since it is always more convenient to print it upon the left-hand side of the diagram, so that anyone approaching it knows before hand what the scale is. The left-hand side scale is more logical in every way, since we always read from left to right, though when the diagram is a wide one it is an aid to reference to give also a scale on the right-hand side. The dividing lines between the various periods of time are indistinguishable in the lower part of the diagram, and this is also a disadvantage. This particular point may be overcome by showing the dividing lines in white, as has been done on the left-hand side of the diagram; but this entails more work, and from a commercial point of view the diagram must be capable of speedy

as well as easy preparation. Such a diagram as this does not lend itself to easy comparison, nor can any other factors be shown on the same chart in order that they may be compared with the main factor.

(b) The " Bar " Diagram.

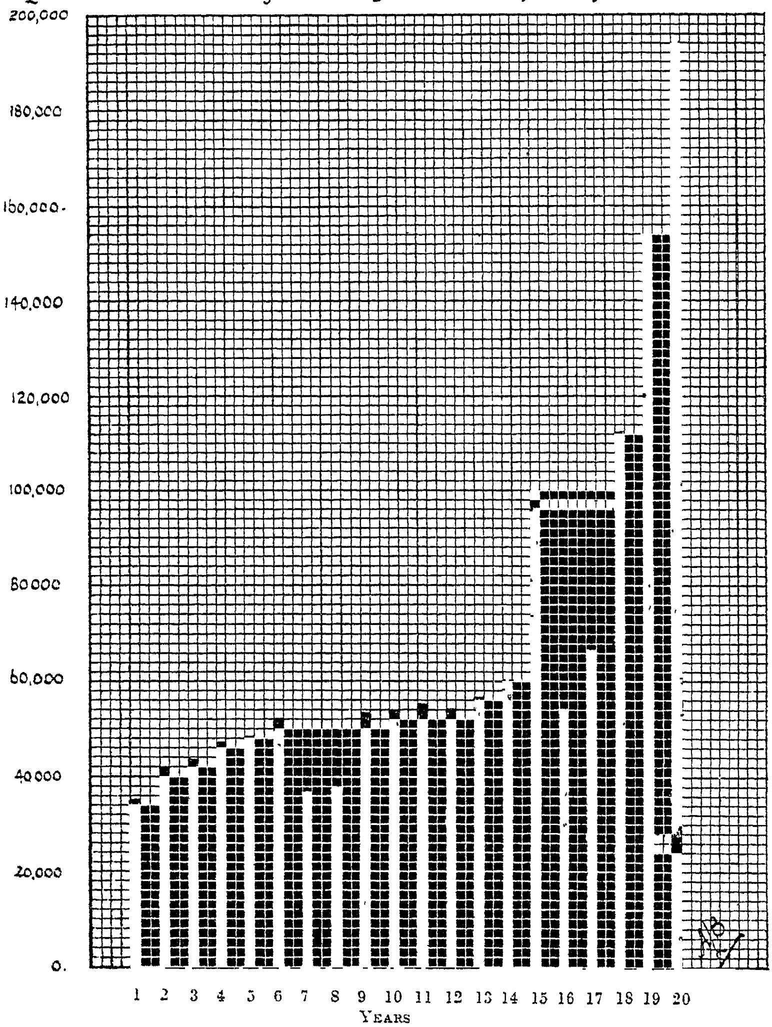
In Diagram No. 7 we see the same data plotted in such a manner as to show each year's Turnover as a separate bar. A comparison of this diagram with No. 6 will at once show that it is not only clearer and more distinct, but that the magnitude of the Turnover is more easily ascertainable from the scale, while the relationship between the Turnover of the various years is more apparent, and as the whole diagram is more compact, it is more useful. This type of diagram, which is prepared on similar lines to No. 1, is one which is easily understandable by anyone to whom a Table of figures would not appeal. The fluctuations, however small, are clearly shown, and the general result is more attractive than that presented in the previous diagram, but it has the disadvantage also that only one factor can appear at one time ; thus when comparison becomes necessary additional charts have to be prepared, and then laid side by side. It would, for instance, be impossible to insert the Moving Average, and so eliminate the short-term fluctuations in order to obtain the Trend.

(c) The " Graph."

It follows, therefore, that for all practical purposes we are compelled to use the Curve, since this enables us to obtain the best results, and, moreover, allows space for other curves to appear on the same diagram for comparison. Whichever of the three methods shown is used, the initial work of plotting is the same, hence when

DIAGRAM №7.

Diagram showing Annual Turnover for Twenty Years.



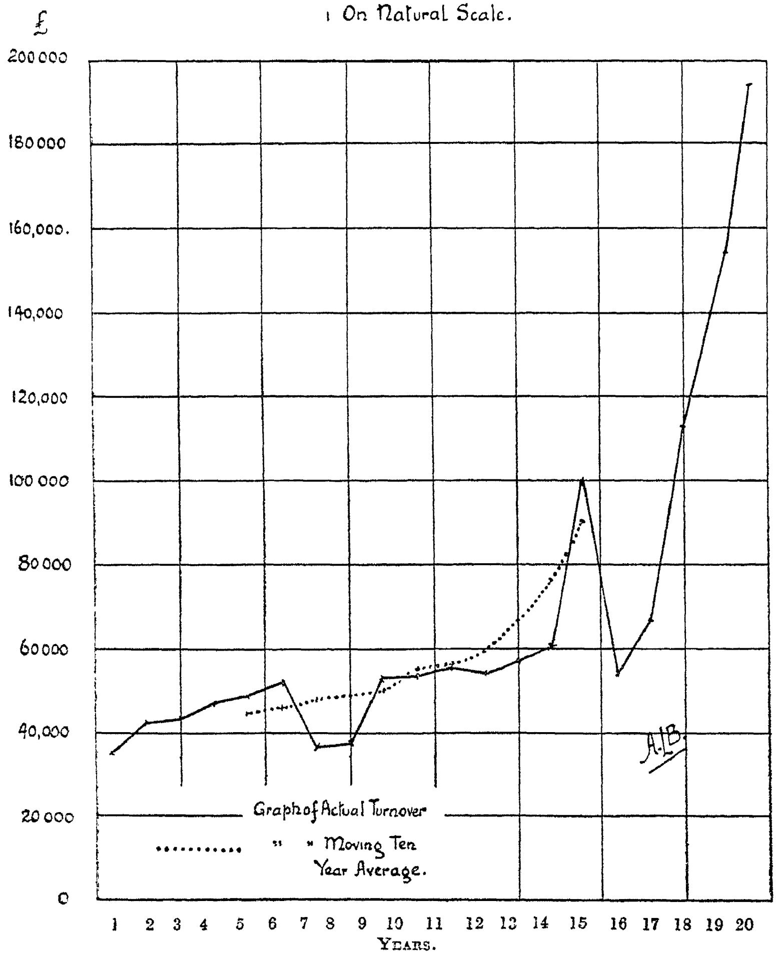
this has been done it is as easy to draw a curve as to prepare Bars or fill in rectangles. In Diagram No. 8 the data appearing in the two previous diagrams have been plotted to the same scale as in No. 7, and the curve has taken the place of the bars. While this form of diagram is, at first, perhaps not quite so easy to understand as the previous one, a little practice will enable even an uneducated person to read it correctly. The rises and falls of the curve tell their own story, and anyone would be able to see *at a glance* exactly what the position was. This rapidity in grasping the significance of the data presented is one of the most important advantages which the graphic method possesses over tabulation, and should prove of great benefit to a busy man desirous of being kept in touch with the details of his business. The time taken in preparing the graph is no more than that spent on a tabulated statement, and, moreover, it can be prepared by anyone, provided that rules are laid down for the purpose of guidance. As a matter of fact, in a business house the Charts would be standardised wherever possible after the utility of the various types had been tested, and from that point onwards the only necessity would be to mark the data on such standardised forms, and draw the curve, and this calls for no great skill. Another important advantage of the curve is that it shows the data as a continuous line, and thus is continuous in its effect, whereas by the Bar method the data for a particular year is shown as a separate entity, which in the case of a business is hardly to be conceived. This continuity of the diagram gives the business man valuable information as to the progress or otherwise of his business.

It will be apparent from Diagram No. 8 that the Curve takes up but little of the space available for the

DIAGRAM No 8.

Diagram showing ANNUAL TURNOVER for TWENTY YEARS.

On Natural Scale.



purposes of the diagram, so that plenty of room remains for the plotting of factors which relate to, or affect the data shown in the principal curve. In the diagram under review the ten-year moving average as shown in column (*d*) of Table F has also been plotted to show the Trend, and it becomes at once apparent that heavy though the various falls have been the elimination or reduction of these short-term fluctuations has had little or no permanent effect upon the progress of the business, for the moving average line moves steadily upward throughout the whole of the period. This upward trend is specially noticeable in the last four years. It should be particularly noted that the moving average is so plotted that each point falls in the middle of the period from the figures of which the average is calculated. This method of plotting an average should always be adopted, for by this means it is possible to discern the periods covered by the computation in an easier manner than would be the case were the points plotted to coincide with the final year of the period.

(1) The Natural Scale Graph.

This method of plotting the actual numbers or results is known as the **Natural Scale Method**, and though it has very many valuable advantages, it suffers from the very serious disadvantage that we are obliged to show rises and falls of equal magnitude by the same vertical distance. Thus, if the turnover in any one year happens to be £2,000, and in the next period £4,000, we have an increase of £2,000, and the curve would move up the vertical scale exactly the same distance as when the turnover increased from £100,000 to £102,000, since the actual increase in each case is the same, though in the first case the increase is 100 per cent., while in the second case it is only 2 per cent.

of the previous year's figures. If the turnover in the second case had increased from £100,000 to £200,000, the line would have risen fifty times the height of the rise shown in the first case, although the ratio of increase was exactly the same, and therefore, unless care is exercised in the reading of the graph, fallacious conclusions as to the progress may be drawn. It is apparent, therefore, that if percentages were plotted we should perceive the ratio which the rises or falls bore to the base year, which would, of course, be equal to 100 per cent. This, again, is a disadvantage, since we frequently desire to know what ratios the fluctuations bear to previous periods, and, moreover, the work of calculating percentages frequently allows errors more or less serious to creep in.

(2) The Logarithmic Scale Graph.

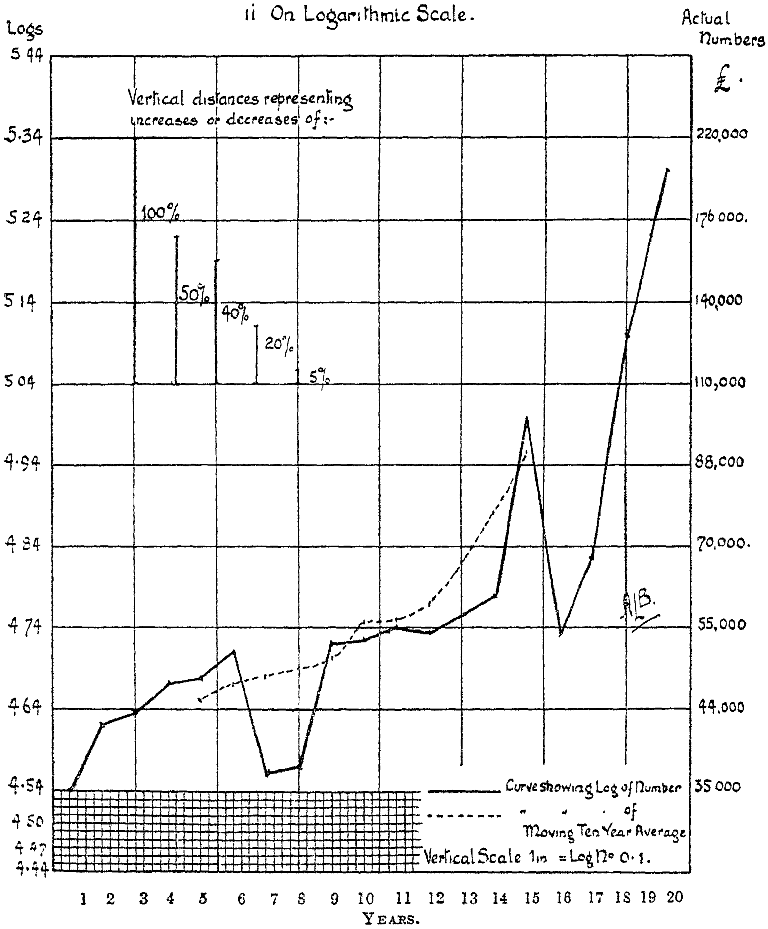
When it becomes advisable or necessary to consider the ratios of increase or fall, some other method of graphic presentment becomes absolutely necessary unless the whole utility of the method is to be destroyed. Such presentation of ratios may, however, be carried out by means of a graph plotted on a **Logarithmic Scale**, and thus the advantages which the graphic method enjoys over the tabulated statement may be continued. If we examine Diagram No. 8 it appears that the increases in turnover which took place in Years 9 and 17 are almost equal in their rises on the vertical scale. As a matter of fact, however, the increase in Year 9 was 41·3 per cent. on the previous year's turnover, whereas the increase in Year 17 was only 23·15 per cent. of the preceding year's figures, a difference that Diagram No. 8 fails to disclose. If we now turn to Diagram No. 9 we see that the difference in the ratio of increase is well defined, the rise in the

curve for the earlier period being greater than that of the latter. This last-mentioned diagram has been constructed on a Logarithmic Scale, on which the Logarithms of the actual numbers have been used instead of the real numbers themselves. As is well known, if we add the Logarithms of two numbers we are actually multiplying the numbers themselves, and by converting the Logarithmic answer to the natural numbers we obtain the product. This is based on the Algebraic formula that when multiplying two expressions we add the indices, thus $x^2 \times x^3 = x^5$, a result obtained by adding the indices of the base x. Logarithms are really the indices of a defined base figure (usually 10), and hence adding the Logarithms of two numbers is equivalent to adding the indices, and thus to multiplying the actual numbers themselves. To construct a Logarithmic Scale we first find the Logarithms of the numbers we desire to plot, and divide our scale into such a number of equal divisions as will allow all the Logarithmic numbers to be included, the Logarithmic numbers progressing in a uniform manner. Thus in Diagram No. 9, one large division on the vertical scale has been made to equal Logarithmic number 0·1, so that each smaller division will be equal to Logarithmic number 0·01. The Logarithms of the actual numbers are then plotted instead of the numbers themselves, with a resultant curve as shown. If we add the Logarithmic Number 0·301 to the Logarithm of any number we double the number. Thus the Logarithm for the number 20 is 1·3010. Add to this the Logarithmic number 0·301, and we get 1·6020, which converted into the natural number gives us 40, or twice the original number, so that in our scale a rise of 3 main divisions will mean an increase of 100 per cent., and other rises in proportion. *The same result is*

Diagram No. 9.

Curve showing ANNUAL TURNOVER for TWENTY YEARS.

ii On Logarithmic Scale.



obtained at any portion of the scale. It must, however, be remembered that in plotting the data on a Logarithmic Scale we are plotting ratios of rise and fall, and not the rises and falls themselves. A careful comparison of Diagrams 8 and 9 will reveal the fact that the fall in the Turnover which occurred in year 7 is emphasised in the latter chart as being more serious than would appear in the former, while the final advances which took place in the 19th and 20th years do not appear so large, owing to the fact that though the actual figures increased considerably the ratios did not move in the same degree. A diagram prepared on a Logarithmic Scale is very useful just so long as it is remembered that the movements of the curves indicate ratios only. If, however, diagrams are prepared on both the Natural and Logarithmic Scales much useful information can be obtained by a comparison of the two. No special skill is required to construct a Logarithmic Scale Curve beyond the elementary knowledge necessary to find the Logarithmic Number of a Natural Number by the aid of one of the numerous Logarithmic Tables obtainable. The value of the Chart may be increased if the natural numbers are shown on a separate scale on the right-hand side of the diagram, as from it an approximation of the actual numbers represented may be obtained, but care must be taken to rely only upon those of the larger divisions, unless conversion be made of the Logarithmic number of any other point for which it is desired to obtain the actual figures of the data. Bars should be shown on the chart to indicate the ratio of increase, and these bars are very useful for measuring the actual ratio of rises and falls in the curve.

Whenever the Natural Scale is being used, the base line of the diagram should always represent zero, and

if this is impracticable by reason of the fact that by so doing the Chart will be too large to be conveniently handled, its position should be indicated, so that no erroneous conclusions will be drawn as a result of the true perspective of the increases and decreases not being apparent. This is a very serious error of many of the diagrams now being presented to the public. In using the Logarithmic Scale, however, no base line should be inserted, or fallacious conclusions may be drawn from the diagram. Another great advantage of the Logarithmic Scale is that the curves may be placed at any part of the diagram which is most convenient, since the ratios are the same at whatever point the curve is plotted, and thus two curves may be brought into close proximity for the purposes of comparison. This advantage is denied the natural scale, for if the scales on which the data is plotted are varied the true relationship is not disclosed.

There are other methods of Graphic Presentation which lend themselves particularly to the display of certain types of data, but great care must be exercised in their use, and they frequently entail more work in preparation than does the graph proper.

(d) The "Bar" Diagram to show Components.

The Bar Method of presentation is a popular one, and is particularly useful when one is desirous of showing the components of which a total is composed. The following Table shows the total amount of cheques "cleared" through the London Banker's Clearing House, *in millions of pounds sterling*, during two successive years, and also how the total is made up, divided into the three clearings, known respectively as the Town, Metropolitan, and Country Clearings.

Year	Town.	Metropolitan.	Country.	Total.
1	£12,329	£1,074	£1,872	£15,275
2	£15,700	£1,177	£2,244	£19,121

In Diagram No. 10 we see these results reproduced in the form of a Bar diagram. This method has much to recommend it, for it is clear, easily understood, and, moreover, shows very definitely the changes which take place in the components of the Total. The Bars should always be constructed so that it is clear that only the linear dimensions are being compared, for if they are made too broad the proportion appears to have been considerably altered, a fact which may be proved by experiment. Any figures which may be appended should appear on the left of the diagram, as in the specimen, as by doing this nothing is allowed to interfere with the varying ends of the bars themselves, for if figures are placed after the bars there is a tendency for the eye to compare the spaces occupied up to the ends of the figures, instead of to the ends of the bars, thus leading to wrong conclusions. When constructing a number of bars for purposes of comparison care must be taken to plot the components in the same order in each case, otherwise the utility of the diagram will be diminished if not actually destroyed. This kind of diagram may be distinguished from the Graphic method proper by describing it as a **Pictogram,** and though it lends itself admirably to comparisons, when only one factor capable of being represented by linear measurements is concerned, care must be exercised if other factors necessitate the use of diagrams introducing area or volume.

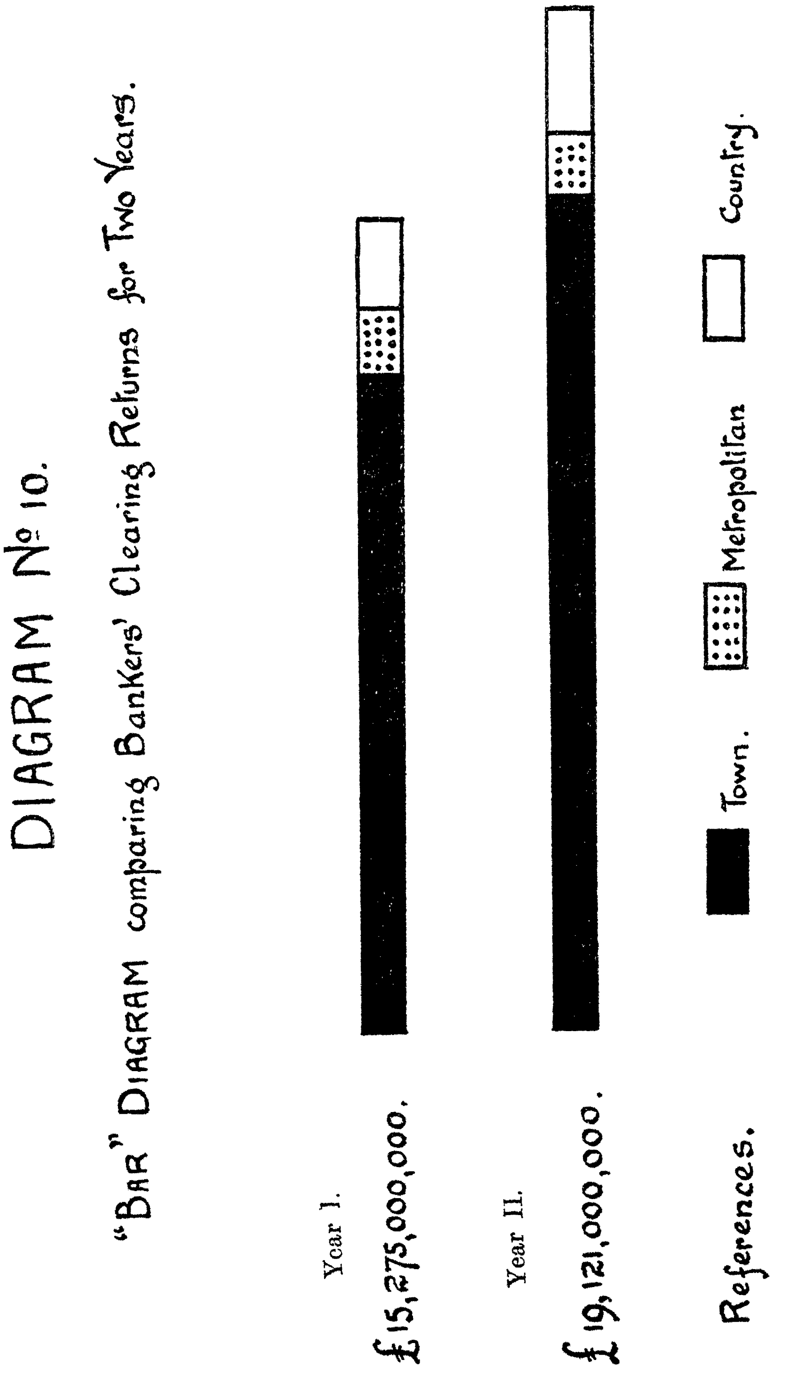


DIAGRAM No. 10.
"BAR" DIAGRAM comparing Bankers' Clearing Returns for Two Years.
Year I.
£15,275,000,000.
Year II.
£19,121,000,000.
References.
Town.
Metropolitan
Country.

(e) Comparisons of Areas and Volumes.

When plotting diagrams involving area, for the purposes of comparison, it must be borne in mind that *the dimensions must vary as the square root of the area*, and not on any other scale. If, for instance, we present certain data in the form of squares, the second set of data being twice the size of the first, then if the area of the first figure be 1 square inch, the second will need to be 2 square inches, and the length of the base necessary to construct the figure can only be obtained by calculating the square root of the area, *i.e.*, $\sqrt{2} = 1{\cdot}415$ inches. If the base of the first were, as it is, 1 inch, it would give a totally incorrect result to plot the second with a base of 2 inches, as by doing this the area of the second square would be 4 square inches, and therefore four times the size of the first. In Diagram No. 11 the difference is clearly seen. Figure (a) is the original, and it is desired to show another figure twice the size of (a). This has been done in figure (b), whereas figure (c) has been constructed with a base twice the size of the first figure. The result of comparing figures (a) and (c) would be to give a fallacious idea of the true proportions. This same difficulty also appears when diagrams constructed in the form of circles are used, since here again we can construct circles for purposes of comparisons on the basis of area, and on the basis of the diameter. Although the former is the more correct method of working, neither method gives very satisfactory results. In Diagram No. 12 figure (b) has been constructed with a diameter twice that of figure (a), while figure (d) has been drawn with an area twice that of (c), which coincides in all respects with figure (a). It appears that figure (b) is far more than twice the size of (a), while figure (d) does not appear to be large

DIAGRAM № 11.

Comparison of Squares on the Basis of Area and Base.

Fig. a.

Fig. b.

Fig. c.

DIAGRAM № 12.

Comparison of Circles on the Basis of Diameter and Area.

Fig. a.

Fig. b.

Fig. c.

Fig. d.

enough to be twice the size of (a). It is this fact that makes circles unsatisfactory wherever comparison is the object of the diagram. They can, however, be used with satisfactory results for diagrams showing the proportions which components bear to the whole. Diagram No. 13 shows such a circular diagram, divided to show how the total of the London Bankers' Clearing Returns for one year is divided. The figures used are the same as in Diagram No. 10, and though it is very clear, and would be improved were the various divisions coloured, it cannot be used like the Bar method for comparison with other years. The best method to be adopted in the construction of such a circular diagram is to make a Table showing the ratios the components bear to one another, and then by ordinary proportional methods divide the number of degrees in a circle in the same proportions.

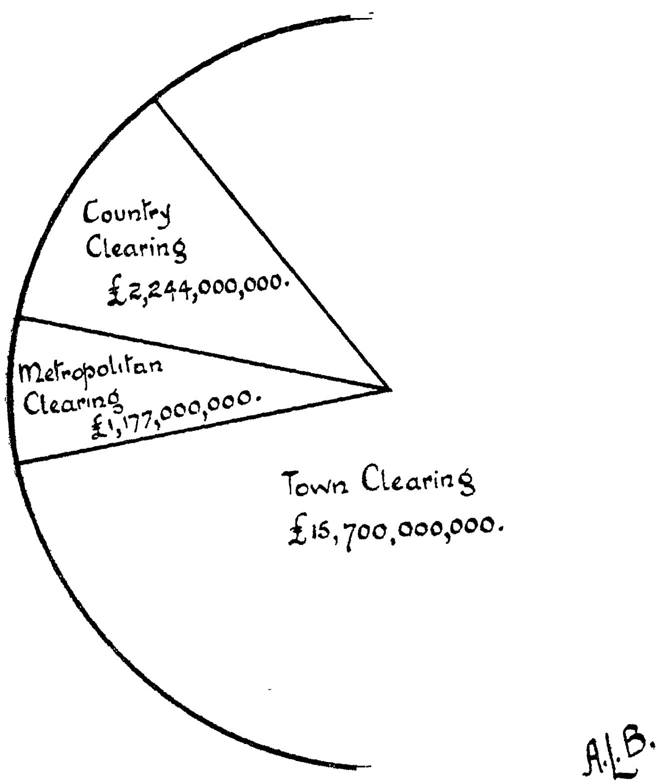
Diagrams showing comparisons of volumes are never satisfactory, for it must be remembered *that the dimensions must vary as the cube root of the volume*, and the adjustment of the various dimensions calls for great care if the proportions are to remain unchanged, since if only one of the measurements be changed the change in the ratio is seldom very apparent.

(f) Diagrams to show Three Factors.

Rectangular figures may occasionally be used to show three factors, such as is done in Diagram No. 14. Here the horizontal scale shows the profit made on each article sold, the vertical scale the number of the articles sold, and the area the amount expended on various items, and also the actual profit realised. In the case of figure (a) the profit per unit is seen to be £5, and therefore, with the standing charges amounting to £60 for Salaries, £20 for Rent, £20 for Expenses, and

DIAGRAM No 13.

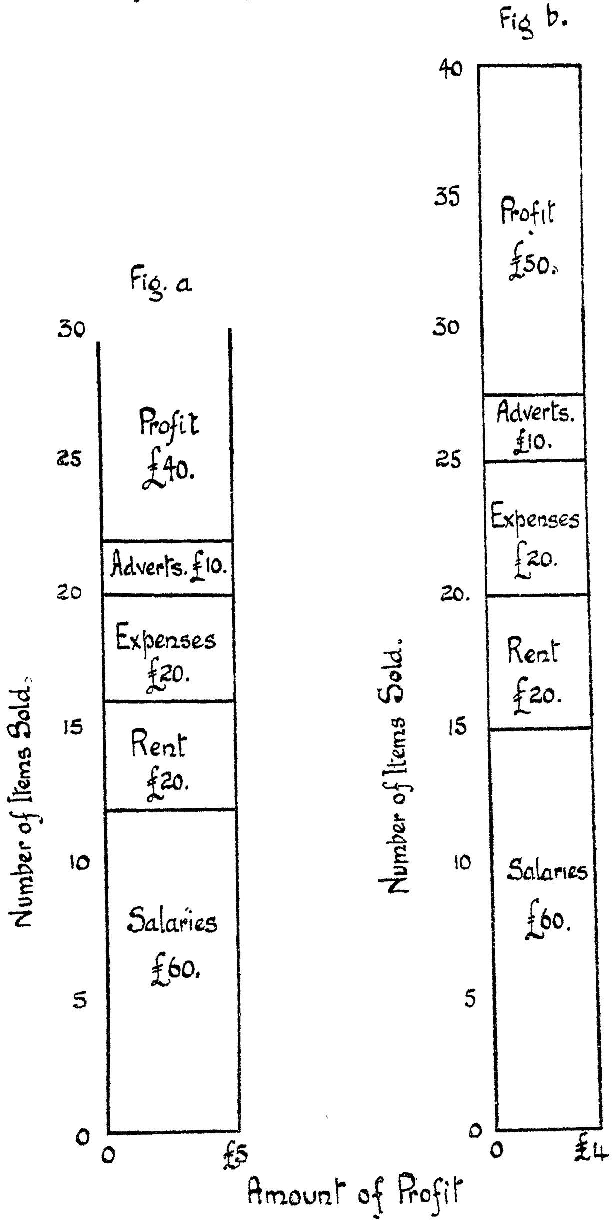
Circular Diagram showing Banker's Clearings for One Year.



TOTAL: £19,121,000,000.

DIAGRAM No 14.

Illustration of Rectangles showing Three Different Factors.



£10 for Advertising, it is obvious that twenty-two units must be sold in order that such standing charges may be covered. Any article sold above this number will yield a clear profit, so that if thirty articles are sold in the period under review a profit of £40 is realised. If this is the capacity of the sales organisation—that is to say, if no more can be sold without an increase in the expenditure—then there would be no object in increasing sales unless the return was more than commensurate with the increased expenditure; but if the organisation is such that additional sales can be made without necessitating further expenditure, then it may pay the business to increase the sales up to the maximum obtainable on the fixed standing charges as shown, and the profit per unit sold might be reduced with the object of creating an extra demand as a result of the lower price. If we assume the Sales capacity to be forty units, then by reducing the profit per unit to £4, it will be necessary for twenty-eight (actually $27\frac{1}{2}$) to be sold before we clear the standing charges; but if we succeed in selling the whole of the allocation of forty units, we shall make a profit ot £50, thus benefiting by the reduction in the profit per unit. Figure (*b*) shows the facts very clearly, and if such diagrams were prepared on the known facts, the management of a business might be able to judge policy more ably, to the advantage of all concerned. This particular method is not suitable for any other type of presentation, and great care must of necessity be exercised in its use.

(g) Diagrams showing Two Scales.

Whenever it becomes necessary to show two scales on the same chart, it is imperative that great care be exercised in the adjustment necessary, to show the true relationship which exists between the two series of

data being plotted and compared. If, for instance, the vertical scale on the one hand represents a greater number than does the other scale, the rises and falls of the curves will vary inversely as the scale, and hence the true perspective is not available for comparison.

§ 6.—Advantages and Disadvantages of the Graphic Method.

Diagrams of all kinds, and particularly the graph, convey strong and vivid impressions to the mind of the reader, and these impressions are not only rapidly obtained, but also retained in a manner not possible from tabulated figures. They tell their story simply, and require a minimum of effort on the part of those to whom they are presented. On the other hand, however, details which are contained in a tabulated statement are not so clearly shown, and frequently considerable trouble has to be taken before such details can be obtained. Diagrams convey strong general impressions, sometimes at the expense of exact details, and hence are only suitable for simple data. If simplicity is adhered to we may sometimes give a mistaken, distorted, or partial statement of the facts. On the other hand, if we give too much attention to detail we may render the diagram so involved that its usefulness is destroyed or materially impaired.

Graphs cannot be quoted in the same way as tabulated statements, and while such tables of figures always convey the same information in the same way, yet the diagram has no standardised value, for if varying scales be used the relationship of the items to one another may and do change materially. This may be clearly seen by plotting the same data on different horizontal scales, the vertical scale remaining unchanged. The narrowing of the horizontal periods will

result in rises being accentuated, and falls would be steeper than on the original scales, while if the space between the horizontal points is increased, rises and falls will both be "flattened out" very materially.

As already stated diagrammatic presentation can only be possible after proper tabulation has taken place, and is extremely useful when it adds simplicity to the presentation of the data. The use of diagrams therefore will depend largely upon the object for which the data is required, and the person to whom it is being presented. When used for the purpose of focussing a large number of operations into a small compass, with the object of conveying information on the general aspect, they are very useful, and would enable a busy head of a business to grasp the essential factors rapidly and without much trouble.

§ 7.—Rules for Constructing Diagrams.

The following rules may be laid down to guide those desirous of obtaining the most satisfactory results from the use of Graphic Method :—

1. The general arrangement of the diagram should be from left to right.
2. Quantities should be represented by linear magnitudes where possible, since areas and volumes are more likely to be misinterpreted.
3. The vertical scale for a curve, wherever practicable, should be so selected that the zero line will appear on the diagram.
4. If the zero line of the vertical scale will not normally appear on the curve diagram, the zero line should be shown by a horizontal

break, or by making the base line of the diagram broken. The former method is, however, preferable.

5. The zero lines of the scales for a curve should be sharply distinguished from the other co-ordinate lines.

6. When curves represent percentages it is desirable to emphasise in some distinctive manner the 100 per cent. line, or other line, used as the basis of comparison, and in such cases the location of the zero line need not be indicated.

7. When the scale of a diagram has reference to dates, and the period represented is not a complete unit, it is better not to emphasise unduly the first and last ordinates, since such a diagram represents neither the beginning nor end of time. When the diagram is brought up to the previous year to that in which the diagram is constructed, the last ordinate should not be ruled in at all.

8. It is better that only those co-ordinate lines necessary to guide the eye when reading the diagram should be shown on the diagram.

9. It is advisable that the curve lines of the diagram should be very sharply ruled, so as to show the curves as distinct from other lines.

10. In curves representing a series of observations it is advisable, whenever possible, to indicate clearly on the diagram all the points representing the separate observations.

11. The horizontal scale for curves should usually read from left to right, and the vertical scale from the bottom to the top.
12. When curves are drawn on logarithmic co-ordinates the limiting lines of the diagram should each be at some power of ten on the logarithmic scale.
13. Figures for the scale of a diagram should be placed at the left and at the bottom, or along the respective axes.
14. It is often desirable to include in the diagram the numerical data or formulæ represented.
15. All lettering and figures on a diagram should be placed so as to be easily read from the base as the bottom, or from the right-hand edge as the bottom.
16. The title of a diagram should be made as clear and complete as possible. Sub-titles or descriptions should be added where necessary, in order to ensure clearness.
17. Where two or more curves are shown in the same diagram, they should be ruled in different coloured inks, or varying types of lines.

SYNOPSIS TO CHAPTER VII.

THE GRAPHIC METHOD APPLIED TO COMMERCE.

§ 1.—Records of Sales.
2.—Utility of the Moving Average Line.
3.—Comparisons of Quantities and Values.
4.—Purchases and Sales.
5.—Price Fluctuations.
6.—Gross Profit and Expenses.
7.—Turnover and Net Profit.
8.—Departmental Records.
9.—Financial and Control Records
10.—Factory Records.
11.—Cost Records.
12.—Transport Records.

CHAPTER VII.

THE GRAPHIC METHOD APPLIED TO COMMERCE.

We have already seen that many advantages accrue from the use of the graphic method, the principal ones being the ease of compilation, and the readiness with which the graphs can be read and the trend and the fluctuations followed. From time to time every business man finds it necessary to call for data relating to such things as Sales, Purchases, Stock, Expenses, Cash Balances, etc., and if these were presented to him in graphic form in such a manner that comparisons could be made between two or more periods, or two or more correlated items, they would do much to save the time of a busy principal and leave him free to devote his time to other matters, instead of having to spend a considerable time in analysing tabular statements. For such a purpose the disadvantages of the graphic method are almost negligible, for it is the broad aspect of matters which needs to be presented in such a connection. Subsidiary curves for the detailed items could also be prepared, and thus enable the Trend displayed therein to be seen and watched. Many businesses would show better returns were data regularly presented to the heads of departments or firms, so that they could see

the exact position, and take such steps as expediency dictates for maintaining sales or output, keeping down expenses and eliminating waste.

§ 1.—Records of Sales.

In a trading concern one of the principal points to be watched is the turnover. The following Table shows the Sales for each month of two successive years, called for the sake of convenience "This Year," and "Last Year," together with the accumulated sales month by month, so that the total turnover up to the end of any month can be obtained without trouble.

TABLE S.

TABLE SHOWING THE VALUE OF MONTHLY AND ACCUMULATED SALES FOR EACH OF TWO YEARS.

Last Year.		Month.	This Year	
Monthly Sales	Accumulated Sales		Monthly Sales.	Accumulated Sales
£	£		£	£
16,200	16,200	January	18,500	18,500
19,800	36,000	February	20,200	38,700
18,500	54,500	March	19,100	57,800
20,200	74,700	April	19,800	77,600
19,200	93,900	May	18,200	95,800
19,500	113,400	June	17,500	113,300
20,000	133,400	July	16,300	129,600
20,200	153,600	August	16,500	146,100
20,200	173,800	September	18,700	164,800
21,000	194,800	October	20,000	184,800
20,100	214,900	November	19,800	204,600
20,800	235,700	December	21,000	225,600

These data have been plotted in graphic form and are reproduced in Diagram No. 15. Clear though the Table may be to a trained mind, the graphic presentation of these figures is striking in the completeness and exactness with which the data are portrayed. In this diagram two different scales have been used, the

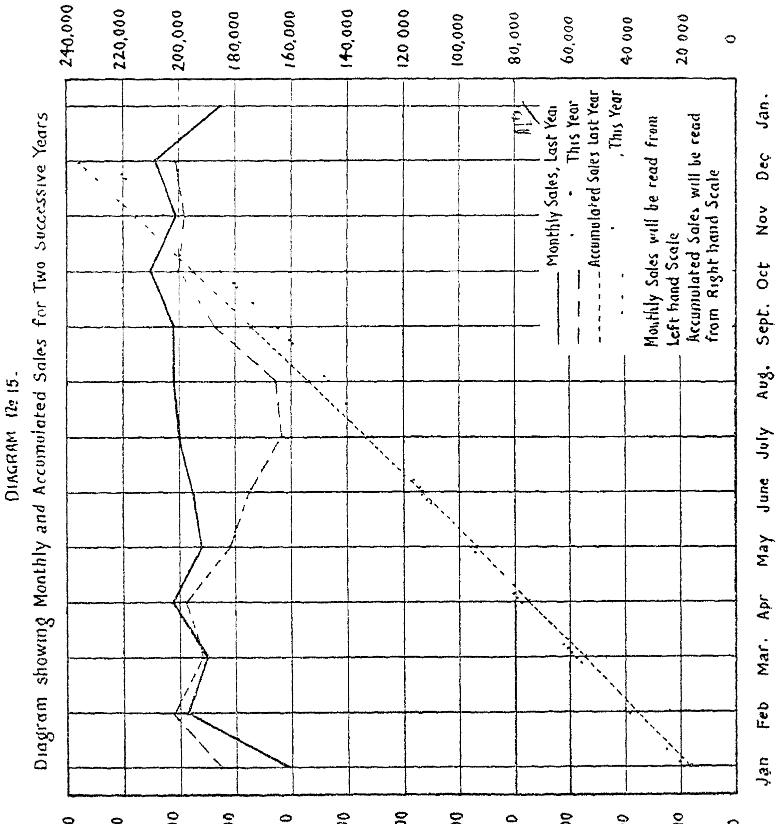


DIAGRAM No 15.
Diagram showing Monthly and Accumulated Sales for Two Successive Years
£
24,000
22,000
20,000
18 000
16,000
14,000
12,000
10,000
8,000
6,000
4,000
2,000
0
240,000
220,000
200,000
180,000
160,000
140,000
120 000
100,000
80,000
60,000
40 000
20 000
0
Jan Feb Mar. Apr May June July Aug. Sept. Oct Nov Dec Jan.
Monthly Sales, Last Year
" " This Year
Accumulated Sales Last Year
" , This Year
Monthly Sales will be read from Left hand Scale
Accumulated Sales will be read from Right hand Scale

left-hand one for use when it is desired to read the curves relating to the monthly sales, and the right-hand one (each space of which is valued at ten times more than the left-hand scale) when references have to be made to the accumulated sales curves. By this means two distinct factors closely related to and dependent upon one another are capable of being read from the one chart, but as the data are distinctive and not correlated the curve for monthly sales cannot be compared with the accumulated sales. As in the case of the Table, the data would be recorded for the current year from month to month as the figures become available, and thus the chart would always be ready for instant reference, the latest available information being shown in its true perspective to the rest of the data. Comparison could thus be made month by month of the figures for the preceding period with those of the corresponding period of the previous year, the causes of divergence inquired into, and, when necessary, steps taken to rectify errors or faulty methods of carrying on the business revealed by such inquiries. The reason, or reasons, for any abnormal fluctuations should always be noted for future reference, as, for instance, where a special type of advertising campaign produces much better results at a certain season than at any other, for if such facts are known, and the knowledge so obtained is applied when similar conditions are again present, sales may be maintained, or even increased, and thus the progress of the business may be assisted. Similarly, if the business be one in a holiday resort, weather may have a very serious effect upon the turnover, and information of this nature is necessary when only "Short Term" fluctuations are being considered, and, moreover, the data will indicate to the head of the business the extent to

which it might be advisable to take out a policy of insurance against the loss which might be occasioned by continuous bad weather. The lines showing the Accumulated Sales are also extremely useful, for if the total sales to date are "lagging" behind those of the previous year, inquiry would be caused and thus steps could be taken where possible to increase sales, as the fall may be due to the goods purchased being unsuitable, insufficient advertising, lack of initiative on the part of the sales staff, price cutting by local or other competitors, or to a general slump in trade. Most of these can be rectified if only the head of the firm is made aware of what is wrong. The best method of presenting the position as far as the business is concerned is by the graph, for every fall is clearly shown, and no time is lost before making the necessary inquiries. Many busy men place tabulated matter aside till a more auspicious occasion for studying it presents itself, and when this examination is postponed the mistakes and faults continue, and the business suffers. The presentation of a graph to the principal attracts his attention to the salient points, and he can thus act without delay, stop leaks, and prevent waste going on to the detriment of the business as a whole. In actual practice each curve on a chart would be drawn in a different coloured ink, thus rendering each clearly distinguishable, though if the lines are not very numerous varying types of lines may be used advantageously, as has, of necessity, to be done in the diagrams used to illustrate the arguments advanced in this manual.

§ 2.—Utility of the Moving Average Line.

It would probably be advisable in a diagram of the nature illustrated in No. 15 to plot and draw a curve

showing the average monthly Sales for, say, a period of five years, or such other period as would cover a complete cycle of trade. Comparison of the current year's line with the curve showing such average would instantly give important information as to whether the progress of the business was being maintained, for, as already has been emphasised, either of two actual years' results may be abnormal, and thus comparison is rendered difficult or impossible for practical purposes, or for purposes of inquiry or reference. In Diagram No. 15 it will be seen that there is a remarkable drop in the sales in "This Year," from May onwards, and this may be due to a customary fall in the sales, or alternatively to the sales of the preceding year being abnormally high. If, however, the Moving Average Line of Monthly Sales were shown and a fall appeared during the same period, the observed result is not unusual, but is due to seasonal or other recurrent conditions. Every business recognises the fact that there are certain periods in every year when Sales fall off for a usually well known reason. In the Motor Trade, for example, it is a well established fact that during October of each year business is stagnant, due in a very large measure to the fact that prospective customers are awaiting the Annual Motor Show, in November, to see the new season's models and their prices before making their decision as to the car they will buy. Similarly, nearly all businesses expect a period of depression immediately after Christmas, and hence endeavour to maintain their returns by the special "Sales," which are a feature of the opening weeks of every year.

Curves to show the same data for other years can also be plotted on this Chart, provided that different coloured lines or distinctive types are used, but care

must be taken not to overburden the diagram, as if this is done its legibility is destroyed, and it becomes an involved mass of lines.

This particular type of diagram lends itself admirably to a large number of variations, as, for instance, when another line is added to show the amount of Sales Returns, or, what is better, since such a curve will normally appear at a much lower place on the vertical scale, a line showing the *Net Sales*. If this be done, the curve will be in close proximity to the line showing the total or gross sales, and thus the fluctuations in the returns can be easily compared with the gross sales. Though returns may be expected in all businesses, and vary greatly, yet if they become very heavy it is a sign that the goods do not suit the customers, or the method of manufacture is faulty, and thus expenses are being incurred unnecessarily, a state of affairs which is not in the best interests of the business. If such data were presented in the form of a tabular statement it would be necessary to calculate the ratio which such returns bear to the gross sales, whereas in a graph the curves for Gross and Net Sales should move parallel to one another, provided the ratio of returns to sales was fairly constant, and a widening of the distance between the lines would give instant warning that the ratio was increasing, and so lead to enquiries. Such a graph would actually reduce the work involved in presenting the information in a proper manner.

Though such charts as that illustrated in Diagram No. 15 show the changes which take place monthly in the turnover, yet weekly, quarterly, half-yearly, or seasonal changes can be shown with equal facility; while the turnover of Departments or Branches, or the sales of individual representatives, may also be plotted in the same way.

From such data carefully compiled over a series of years the business man can find the period of the year when his sales are likely to be heaviest, and thus arrange his purchases and deliveries at such times as the goods are likely to be required. Such a policy tends to make the turnover of the goods handled more rapid ; obviates the necessity for maintaining heavy stocks in warehouses or stockrooms with the consequent risk of depreciation and damage ; and does not result in working capital being locked up in stock which is only demanded at certain well-defined periods. If, in addition to compiling data of the financial aspect only, records were maintained of the types of goods sold, and also the quantities, further useful information could be obtained and utilised to the advantage of the business. A certain large London Store recently advertised that its sales records showed that the sales of a particular commodity were always highest in a certain week in each year, and that therefore during this period in the current year they had made special arrangements for the display and sale of such goods. This is an example of really scientific management, the result being that the business applies itself to meeting just that need of the general public which is of the greatest importance at the moment, to the obvious advantage of the public themselves, and to the benefit of the business in maintaining or increasing the turnover as well as in enhancing its general reputation. By such means as this many of the chances of incorrect buying are eliminated ; losses are thus minimised ; and there is a greater possibility of success being realised. All these data are capable of being shown by the graphic method, and thus the busy head of affairs may be kept in constant contact with the movements and position of business with the expenditure of a modicum of time and trouble, leaving details

to the responsible Heads of Departments, but obtaining such information as will lead to enquiries should matters not be developing in a satisfactory manner.

§ 3.—Comparisons of Quantities and Values.

If Statistical methods are to be completely successful in the commercial world, it is essential that records be kept not only of the monetary values, but also of the quantities of goods handled, for as already shown in Chapter II, when any change takes place in the purchasing power of money (*i.e.*, when there is a general change in the level of prices of commodities), the financial turnover will be affected, and wrong conclusions drawn therefrom. If, for instance, the purchasing power of money declines (*i.e.*, the price of commodities increases) and the value of the turnover of the business remains the same, then the business will not be progressing, since the actual quantity of goods sold will be less. Conversely, any increase in the cash value of the sales will not necessarily mean that the turnover of goods has increased. In order to ascertain the true position it will be necessary to compare not only the amount received or charged in respect of the sales, but also the quantities sold, as without this latter information it will be impossible to say that the business has progressed, for though the figures in the financial statements may show an increase, the actual quantities sold may be less than for the previous year; and, after all, the true test of a business is the quantity of goods sold, since any variation in price is bound to affect the financial accounts. A rough idea of the position can be gauged from a study of such figures, allied with the movements which have taken place in the Index Numbers of Wholesale Prices, for, as shown in the Chapter on Index Numbers, these are

designed to give information relative to changes which take place in the purchasing power of money. In Diagram No. 16 an endeavour has been made to bring the two factors of value and quantity together in one chart for a period of ten years. The following Table gives the Turnover in £'s sterling and in Tons:—

TABLE T.

TABLE SHOWING TURNOVER IN £'S STERLING AND IN TONS FOR TEN YEARS

Year.	Turnover.	
	£	Tons.
1	143,400	202,000
2	153,500	202,000
3	200,200	220,000
4	292,500	225,000
5	245,300	219,000
6	224,700	227,000
7	211,600	230,000
8	199,500	190,000
9	295,000	236,000
10 . ..	308,400	241,000

In the diagram the monetary values will be read from the left-hand scale, and the quantities from that on the right-hand side of the chart. From the curves thus plotted much valuable information can be obtained. For instance, the monetary value of the sales increased in year 2, although the actual quantity of goods sold remained the same as in the previous year, while in year 4 the sales only increased from 220,000 tons to 225,000 tons, but the value rose from £200,200 to £292,500. In the first of these cases we should have assumed that the business was progressing if we only considered the financial figures, while in the second we should get an entirely erroneous idea of the progress of the business were the quantities also not available for comparison. In each case the increase in the value of the goods sold is due to the

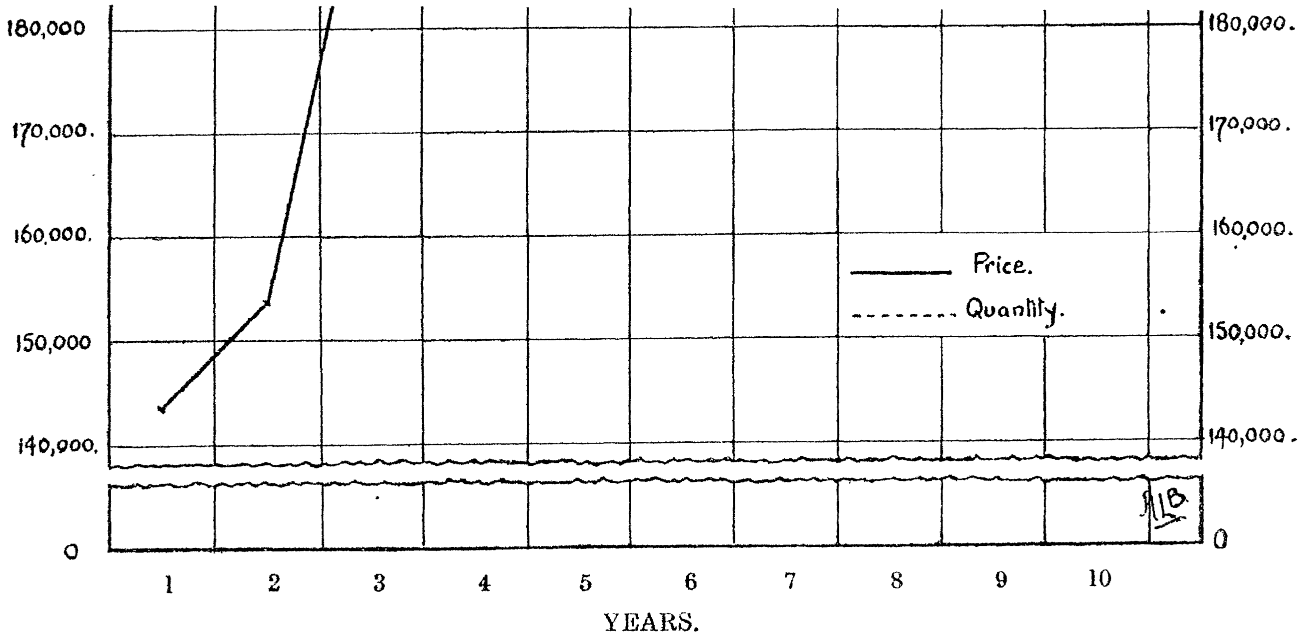
fact that there has been a change in the price of the commodity. If the price were steady, we should expect to see the two curves fluctuating together. It will be seen from the diagram that the quantities sold tended to increase, and that in the ten years there was only one period which records a fall, *i.e.*. year 8; but if we study the financial records we see that falls took place in years 5, 6, 7 and 8.

Great care is necessary in plotting such a chart combining two variables, for if a careful ratio is not calculated in connecting the two scales the real relationship between the two factors is destroyed or damaged. The unit of quantity should be plotted on a scale compiled after very careful study of the ratios existing between the unit of quantity and the monetary value. If, for instance, the records show that the price per unit of the given commodity averages, say, £5, then one unit of quantity and £5 can be taken as the basis of the scales, and each would be represented by the same vertical distance. If some such method be adopted the curves will be closely related, and hence comparison will be just, and the conclusions drawn therefrom will be reliable.

§ 4.—Purchases and Sales.

The relationship which exists between Purchases and Sales is very important, but it is a matter frequently overlooked by business men, with the result that a business often becomes overstocked; valuable working capital is locked up and thus is not available for more remunerative employment in some other department of the business. If the purchases and sales of a business be plotted periodically on the lines shown in Diagram No. 17, it can at once be seen whether the purchases are too heavy, resulting in

DIAGRAM No. 16.



facing p. 196

stocks accumulating unnecessarily, or whether they are not heavy enough, in which case the purchases will be insufficient to maintain the stock at that amount which experience shows to be necessary for the well-being of the business. The following figures are those which have been plotted in Diagram No. 17 :—

TABLE U.

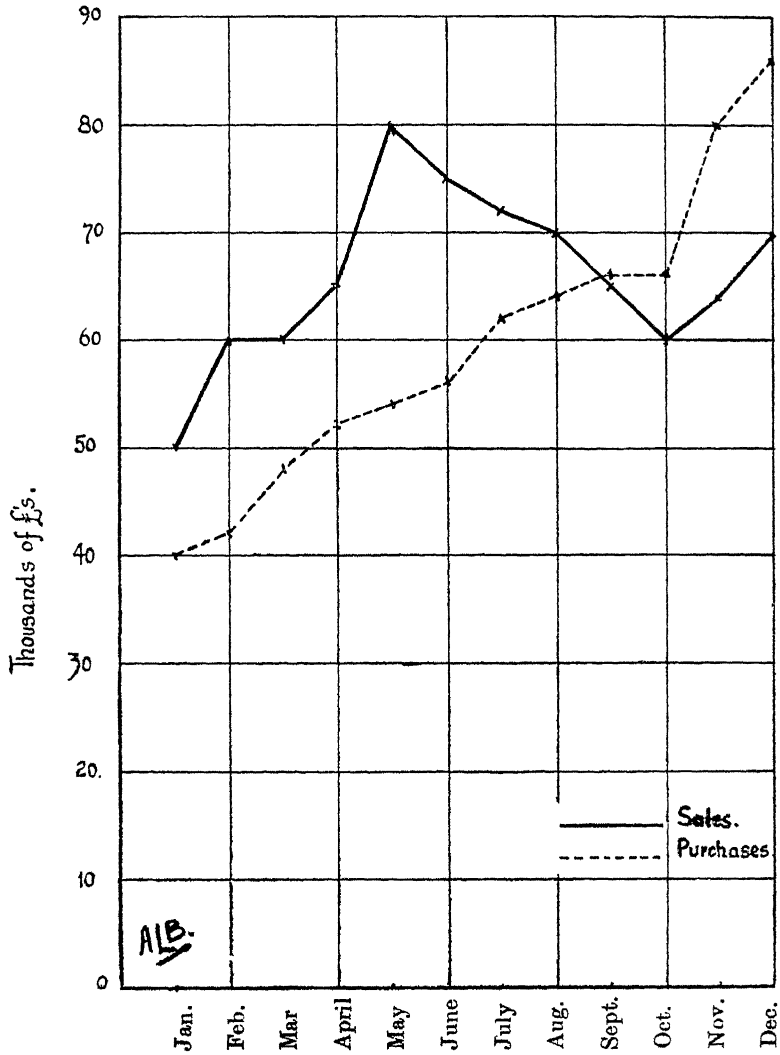
TABLE SHOWING MONTHLY PURCHASES AND SALES FOR ONE YEAR.

Month.	Purchases.	Sales.
	£	£
January	40,000	50,000
February	42,000	60,000
March	48,000	60,000
April	52,000	65,000
May	54,000	80,000
June	56,000	75,000
July	62,000	72,000
August	64,000	70,000
September	66,000	65,000
October	66,000	60,000
November	80,000	64,000
December	86,000	70,000

The diagram shows at once that up to May the purchases were much smaller than the sales, so that either old stock is being disposed of, or alternatively the stock was being gradually depleted. After the month of May the sales fall off, but the purchases line continues to advance, until in September more goods are being purchased than are being sold, that is, of course, from the point of view of monetary value, so that expenditure or liabilities in respect of this one item alone will be in excess of the Cash Sales receipts or amounts charged to Customers' Accounts. It needs but little demonstration to make it clear that unless such movements as these are accompanied by a decrease in the

DIAGRAM No 17.

Chart showing Combined Purchases and Sales for Twelve Months.



amounts owing by customers the cash position of the business will speedily prove unsatisfactory. It may be argued that this is necessary to replenish stocks for the next period, which apparently is the period of rising sales, but it would appear to be poor policy to be buying steadily for eight months in the year in order to sell in the other four, though in certain types of business this procedure may be necessary.

§ 5.—Price Fluctuations.

If records of Sales are made over a number of years an average can be obtained for different periods, *e.g.*, weekly, monthly, quarterly, or by seasons; which if plotted for reference will give an interesting and reliable clue to the fluctuations of this item. Most businesses show a great many variations in their returns, many of the fluctuations in, say, Sales, being well defined, and the periods in which they occur are regular or seasonal in their recurrence, and the average of a number of years will, as we have already seen, eliminate abnormalities, thus giving a base upon which the quantities needed for any period can be estimated, and the necessary purchases made to meet the demand when it arises. While it is true that the period in which it is expected that the goods purchased will be disposed of may be, and sometimes is, abnormal, it is better to base estimates and operations upon known past experience, and so eliminate some, at least, of the factors of chance which are present in all businesses, and which are uncontrollable. It would be possible to guard against the purchases being too heavy, or too small, though we cannot entirely eliminate all differences. If similar data be compiled as to the prices paid for the goods in which

one deals, and the average over a series of years be taken, we shall find that with many of the commodities of everyday life prices tend to be higher at one season than at another. If we know exactly when lower prices are to be expected, arrangements can be made to purchase supplies at such times with a consequent reduction in the cost, thus ensuring greater profit; or, alternatively, placing the buyer in a better position than his competitors who buy their supplies at times when the price is not so favourable. Again, the graphic method is by far the most suitable for presenting the facts of the case. In Table V (*see next page*) we see the average price of English Wheat calculated for a period of ten years. This average has been plotted for the sake of clearness in Diagram No. 18, and those periods when it is better to buy supplies are very clearly seen, so that an operator in wheat who could purchase his supplies in January would be enabled to make quite a good profit, if he were able to obtain such a quantity as would enable him to meet his customers' demands for some months. The best periods in which to purchase wheat will be seen to be January, September and October, or thereabouts, and knowing this fact an operator could make the necessary financial arrangements to meet his needs. Importers and Exporters could compile similar records for the Rates of Exchange, and thus possibly be able to arrange their transactions so that payment would be made or received at such times as the exchanges were favourable to them. In this way goods could be purchased at a lower expenditure, and the purchaser be in a much better position than a competitor who did not pay any attention to the regularity with which many of the exchanges move in normal times.

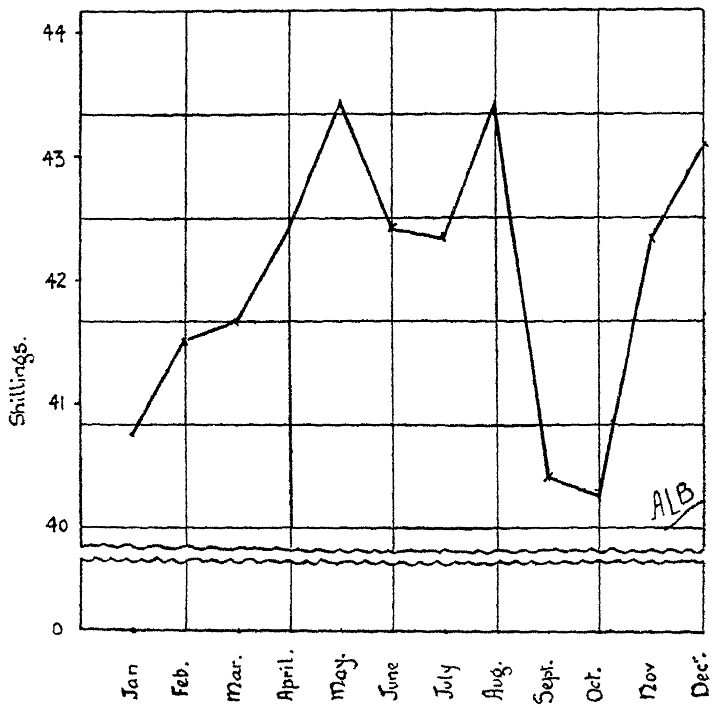
TABLE V.

AVERAGE PRICE OF BRITISH WHEAT PER IMPERIAL QUARTER IN ENGLAND AND WALES IN EACH MONTH FOR TEN YEARS.

Year.	January.		February.		March.		April.		May.		June.		July		August.		September.		October		November		December.	
	s.	d.	s.	d.	s	d	s.	d.	s.	d.	s.	d.	s.	d.	s.	d.	s.	d.	s.	d.	s.	d.	s.	d.
1	35	3	33	1	31	1	31	0	32	9	32	3	30	10	31	4	31	3	31	1	31	6	32	6
2	32	8	33	10	35	6	38	5	41	11	42	9	43	3	42	4	34	7	31	8	32	7	33	2
3	33	6	33	1	32	7	33	4	31	11	29	4	31	3	33	2	31	3	30	2	29	11	30	6
4	30	8	30	4	30	1	30	5	32	0	32	2	32	2	31	8	32	0	32	9	33	2	32	10
5	33	3	34	1	34	1	36	1	37	8	37	2	38	6	37	2	33	2	32	2	32	7	30	6
6	30	6	30	11	31	1	31	5	32	7	32	8	33	5	33	8	31	8	30	11	30	3	31	1
7	31	0	31	0	31	5	31	6	32	7	34	1	34	1	36	9	37	6	37	0	40	3	42	6
8	48	8	54	11	54	7	55	3	60	10	57	5	52	2	54	2	43	6	46	4	52	11	53	9
9	56	10	57	8	56	11	54	1	55	1	49	4	50	0	57	2	59	4	60	7	69	5	73	0
10	75	9	76	2	79	2	83	10	77	10	78	1	78	2	77	11	71	0	70	9	70	3	70	9
Average	**40**	**9**	**41**	**6**	**41**	**8**	**42**	**5**	**43**	**5**	**42**	**5**	**42**	**4**	**43**	**5**	**40**	**5**	**40**	**3**	**42**	**3**	**43**	**1**

DIAGRAM No 18.

Graph showing Ten Year Monthly Average Price of English Wheat per Imperial Quarter.



Clare in his manual on Foreign Exchanges shows by means of diagrams that when the rates of the exchange with America were averaged in this fashion the movements were very regular in their recurrence, and the same is at any rate true under normal conditions with all the gold exchanges, so that there is no reason why the ordinary business man should not reap some of the advantages to be obtained from the knowledge of the possibilities of these movements. It is true that greater interest has been taken recently in the movements of the exchanges, but when these again become normal the possibility of making additional profit by using the exchange fluctuations to the advantage of the business will probably be overlooked. Any " averages " so used would be adjusted from time to time, so that any tendency for the fluctuations to shift would be brought into account. The Moving Average would prove most useful in this connection.

§ 6.—Gross Profit and Expenses.

There are many factors in Commerce which are so closely related to one another that consideration of any of them is best made when the others are also available for comparison and criticism. A good example of this is Gross Profit and Expenses. If we consider expenses only, then any increase may be the subject of needless criticism, since it may be discovered that the gross profit has increased also, and, since the rate of gross profit is not likely to vary, there must have been a greater monetary turnover, which may easily lead to greater expenditure ; so that if it can be shown that the ratio of increase in

gross profit is equal to, or greater than, the increase in expenditure, there will be little or no need for the expenditure to be adversely criticised. It must not, however, be assumed that there is no need to go into the question of expenditure, as this is a necessary procedure from time to time if real economy is to be effected; but if the standard of expenses is exceeded only by a lower ratio than that by which gross profit has been increased, then it is more than possible that the increase in expenses is due to the increase in the turnover, and to that only. In this connection a detailed study of the business is advisable in order that the limits of the economic law of increasing returns may be ascertained, and efforts made to obtain the maximum output or turnover of which the business is capable without further expenditure. If, however, the expenditure has increased and there has been no corresponding increase in gross profit, then it would appear that inquiry is necessary if the increase is to be justified, though it must always be borne in mind that the increase in sales may not always coincide with the increase in expenditure. This would happen when a special advertising campaign has been inaugurated, such as is necessary to place a new commodity upon the market, the full or complete effects of which are not immediately felt, but which bear fruit later when the product has become well known as the result of the campaign, and therefore full inquiry needs to be made before conclusions can be drawn. It would be very useful to the business man to compare Gross Profit and Expenses and also the Net Profit, and this has been done in Diagram No. 19, which is plotted from the following data:—

TABLE W.

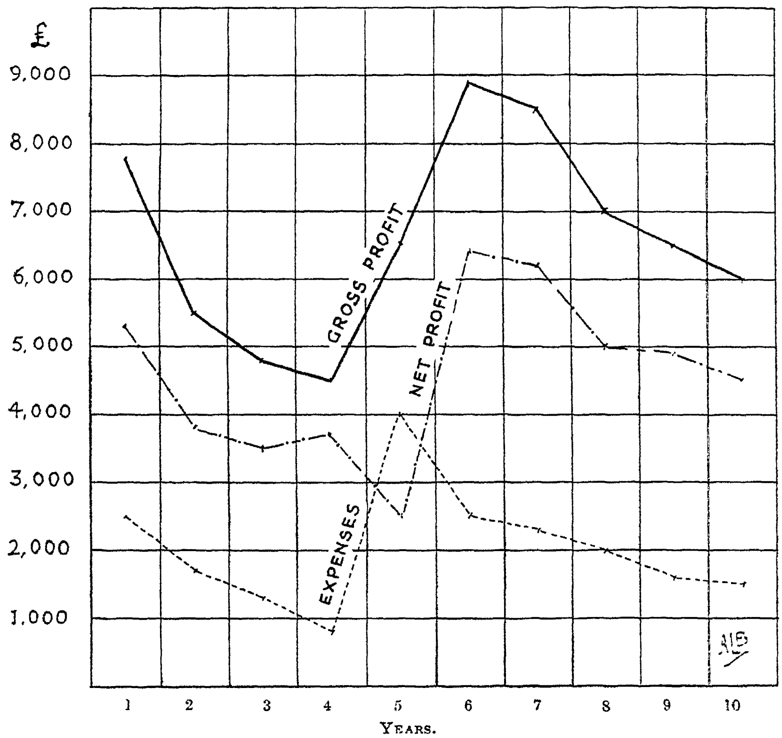
TABLE SHOWING EXPENSES, GROSS AND NET PROFIT OF A BUSINESS FOR TEN YEARS.

Year	Gross Profit	Expenses.	Net Profit
	£	£	£
1 ..	7,800	2,500	5,300
2	5,500	1,700	3,800
3 .. .	4,800	1,300	3,500
4	4,500	800	3,700
5	6,500	4,000	2,500
6	8,900	2,500	6,400
7	8,500	2,300	6,200
8	7,000	2,000	5,000
9	6,500	1,600	4,900
10	6,000	1,500	4,500

This statement presented to a business man would need careful study and analysis in order that the full significance of the figures could be appreciated, but the diagram shows immediately the divergence which exists. From the curves there plotted we can see that in the first three years the gross profit was falling, but the expenses were also falling, and consequently the net profit did not change to the same extent as the gross profit. This is clearly shown by the curves for the two factors approaching nearer to one another. This fall in the gross profit was in all probability due to a fall in the sales, as a reduction in the rate of gross profit is very unlikely, the percentage which is added to cost to cover general overhead charges and profit being a fairly constant factor. It is observable also that in the fourth year the fall in the expenditure was accompanied by an even greater fall in the gross profit, and this constant falling in the gross profit earned while expenses are also on the downward grade would seem to point to the fact that economy in management was so great that it was seriously handicapping the business. This assumption is also borne out by the fact that in the fifth year a sharp rise in the

DIAGRAM No. 19

Diagram showing Gross Profit, Expenses and Net Profit for Ten Years



amount of expenses was accompanied by a considerable advance in the amount of the gross profit, though the advance in the latter case was not of the same ratio to the previous year as the expenditure, with the result that the net profit in that year actually fell by £1,200. The advance in the gross profit was, however, even greater in the next year, though expenses actually fell considerably over those of the previous year, and from this it would appear that in the fifth year some special effort was made to produce better results from the business, and the effect of this effort was also apparent in the following year. Subsequently the gross profit again declined, and this decrease was even more sharply apparent than the decline in the expenses. If one were investigating the business with a view to improving the results to be obtained therefrom, it would be necessary to call for explanations of the very heavy increase in the expenditure incurred in the fifth year, and it would be wise to analyse the expenses with a view to discovering the effect of each item upon the business. If it appeared that this heavy expenditure was due to an advertising campaign carried out during that year or some part of it, the whole of the cost of which was charged in that year's expenses, we should at once see that the decline in the gross and net profit was due to rigid economy adversely affecting the prospects of the concern. We could then arrange that the "cheese-paring" policy should cease, and the business be placed upon a progressive footing by judicious advertising or other means. Here is an example of where the science of Statistics properly applied can assist a business, and lead to greater profits and expansion. Just as it is possible to be too lavish in expenditure, it is also possible to be too niggardly. By utilising Statistics in such a manner

we are making it possible to trace faults, stop leaks, prevent waste, reduce the ratio of expenditure to turnover; or steps can be taken to render the expenditure more remunerative—in short, to render our management more scientific, and so provide for contingencies whether possible or probable, thus eliminating or reducing the elements of chance which are present in all business or commercial undertakings.

§ 7.—Turnover and Net Profit.

Where articles are being handled which are uniform in price and quality, the gross profit per article can easily be ascertained, and a chart prepared to show the Gross Profit on Sales, the Standing Charges (including such items as Depreciation, Debenture Interest, and similar expenditure, which have to be met periodically), and the Net Profit made and therefore available for the payment of dividends, etc., so that it will be possible to see at a glance if the sales are sufficient to yield an adequate return. Such a Chart is seen in Diagram No. 20, which has been prepared from the information given in Table X. Each article sold shows a gross profit of £1.

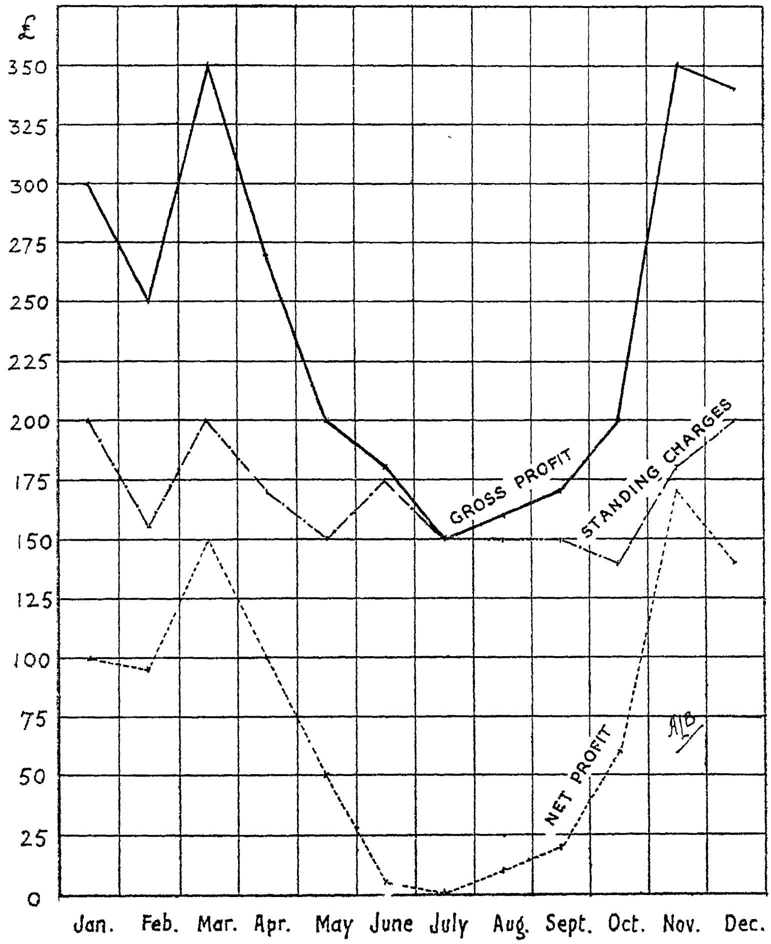
TABLE X

TABLE SHOWING NUMBER OF ARTICLES SOLD, GROSS PROFIT, STANDING CHARGES AND NET PROFIT.

Month.	No. of Articles Sold.	Gross Profit.	Standing Charges.	Net Profit.
		£	£	£
January	300	300	200	100
February	250	250	155	95
March	350	350	200	150
April	270	270	170	100
May	200	200	150	50
June	180	180	175	5
July	150	150	150	0
August..	160	160	150	10
September	170	170	150	20
October	200	200	140	60
November	350	350	180	170
December	340	340	200	140

DIAGRAM Nº 20.

Diagram showing Monthly Gross Profit, Standing Charges and Net Profit



It will be seen from the diagram that the heavy fall in the gross profit due to the fall in the sales for the period subsequent to March was not accompanied by a corresponding fall in the standing charges, and consequently the net profit fell, till in June there was a margin of £5 only, while in the next month the income from sales was only just sufficient to meet the current allocation of expenditure on Standing Charges. Such a state of affairs might conceivably overtake any business, particularly one of a seasonal nature, and if the facts are known, then provision can be made for curtailment of such expenditure as is unnecessary, or arrangements made to meet interest charges at some other time, or credit obtained to carry the business over the slack season. If this sequence of events were regularly recurrent, purchases could be made, so that payment fell due during the periods January to March, and October to December.

§ 8.—Departmental Records.

There is no need for the graphic presentation of facts to be confined to the use of the office staff or the management. The efforts of the sales staff or any organisation might be spurred onward were they in a position to see how their departments or branches stand in relation to the remaining units of the organisation. If suitable Charts were prepared and exhibited to them showing the Departmental or Branch Sales day by day, week by week, or month by month, together with a curve showing previous results, or the results of some similar department or branch, they would be encouraged in their efforts to keep the line for current Sales from falling, thus assisting the development of *esprit de corps* and team work. One large

London Store has already adopted some such idea, and rewards that department which shows the greatest increase in its Turnover. Graphic methods of portraying the positions of the various departments are used, and the scheme has met with considerable success. By showing the various items of expense connected with the department or branch, and the way in which these compare with others, is also a good plan, as the tendency to waste is checked, and it would frequently pay the employers to give a special bonus to the staff of any department showing a reduction in such items as Stationery, Wrapping and Packing Paper, etc.

§ 9.—Financial and Control Records.

Charts could be used to show the Cash and Credit Sales, and thus make adequate arrangements for safeguarding the business against possible loss from Bad Debts owing to an increase in the Credit Accounts. Every business needs to watch this ratio of Cash and Credit Sales to the Total Sales very carefully, especially when the Cash Resources are limited, or the policy of the business is to maintain but small Cash Balances, and depend upon the Cash Sales and the collection of Debtors' Accounts to provide sufficient funds to meet maturing liabilities. In such cases a diminution of the ratio which the Cash Sales bears to the Total Sales may result in the current receipts being insufficient to meet the current expenditure. An extension of the credit accounts either in number or amount means delay in replenishing the cash resources, and this delay may result in very marked financial stringency. Receipts and Payments of cash might also be so recorded, and thus after a series of observations had been made it would be seen at what time of the year

the greatest cash balance was accumulated, a very useful thing to know when the question of adequately financing a business is being considered. A very useful diagram for purposes of financial control is one showing the Accounts Receivable, and the Amounts actually received day by day. Such a diagram will at once show what "lag" exists in the payments by the Debtors of the business. This information is important when a business is operating on a narrow cash margin, and hence would enable the credit terms to be varied in order to give the necessary margin for security. "**Bar**" diagrams showing the active Liabilities, and the Liquid Assets such as Cash, Bills Receivable, and Book Debts should also be prepared periodically, such a diagram giving a picture of the actual financial position of the business.

Various forms of expenditure can be advantageously compared with other factors, and will afford much valuable information. Many large firms spend considerable sums on advertising, and the cost of each of the various forms used should be graphically shown together with the results obtained. From such diagrams it will be possible to ascertain the best form of advertising, or the best media; while, if the time element be also taken into account, it may be possible to find the best days or periods on which to advertise. It may be argued that many of the results will not be capable of being allocated to any particular advertisement, but this can be overcome by spreading such unidentified orders over the various media in the ratio which the known results bear to one another. It must not, however, be overlooked that this enquiry should also be conducted by areas, since the form of advertising which proves most successful in one district may be a failure in another.

The cost of advertisements should be spread over the orders received, in order that the value of the advertisement can be gauged in the cost per unit of sales. Comparative diagrams will show which form of advertising, or which media, is the most remunerative. It may then be found that a more expensive advertisement results in a lower advertising cost per unit of sales. Another diagram could show the Inquiries received, and the orders resulting therefrom, as well as the Total orders received through the advertisement, and so ascertain what proportion of the enquiries resulted in business. The expenses incidental to the actual sale of the goods is also a very important matter, and comparative charts showing Sales in each area, and the direct expenses incidental to the Sales, such as Salaries, Expenses, etc., would give important information relative to the cost of transacting business in the various districts, as well as the cost per unit of Sales of each Salesman or Branch.

§ 10.—Factory Records.

The graphic method can be used to advantage in the factory, and in many ways it lends itself admirably to the presentation of important data to the workpeople, who would probably not take any interest in tabulated statements. One factory producing a large number of similar articles of varying sizes, charts the output of every machine, as shown by an attached automatic numbering machine, and places the chart in close proximity to the worker, so that he can see how his output is progressing day by day, and also compare it with the average output of all the similar machines employed on the same class of work for the previous year, which is also shown on the chart. The total output of each size of article is also plotted daily

on a large chart in the office of the managing director, so that he is enabled to see at a glance whether the output is steady, increasing or decreasing, and ascertain the reasons for fluctuations. Any unusual circumstance is plainly shown on the chart, such as when a machine is not working, etc., and each type of article has its own distinctive coloured line. It has been found that the operatives take a keen interest in their charts, and endeavour to maintain a steady output, and therefore instantly report the slightest defect in their machines which might lead to any reduction of their output. The experience of the last few years shows that since the installation of this system the days lost by breakdowns to machines have almost disappeared, and the output has consequently increased from this source alone. The management can, by comparing the charts for various machines, which is done from time to time, discover those machines which are "lagging" behind the others, and ascertain the reasons therefor. Improvements have been made in the running of several machines as a result of this investigation. They are also enabled to gauge the relative merits of their operatives, and can reward the diligent and weed out the undesirables. The result is so marked and the time taken in so recording the output so small, that it is a recognised part of the firm's organisation. The same firm also maintains charts showing the weekly percentage of time occupied by their employees in non-productive labour (such as cleaning machines and preparing them for the work in hand), and they can thus instantly see whether this item is tending to grow, and ascertain the reasons therefor. This chart has been found to be extremely useful, particularly when the Machine Hour-Rate used in their costing system is under review or revision.

Another chart which could be used in a Producing Business would be one showing the Selling and Cost Price per article or per unit of production, whatever that might be. This would be particularly useful in cases where the selling price fluctuates considerably, and is fixed by the operation of competition, for if the selling price does not show a sufficiently large margin over the cost, then the question of using the plant for some more remunerative work may well be considered.

Among other diagrams which might be used with advantage in Factory Records are those showing the following information :—

(*a*) The ratio the Stocks of Raw Material in hand bears to the quantities used by the Factory.

(*b*) The relationship between Raw Material Stocks and Finished Goods on hand.

(*c*) Comparison of Labour Charges, Cost of Material used, and Standing Manufacturing Charges.

(*d*) The divergence between the time estimated to be taken in completing a job, and the time actually taken.

(*e*) The ratio of spoiled products to total output, together with a comparison with the standard spoilage allowed.

§ 11.—Cost Records.

The Prime, Factory and Total Costs might be shown in one diagram with advantage. The Prime Cost of an article is the cost of the materials used plus the actual cost of the direct labour expended upon the material. The Factory Cost is obtained by adding to the Prime Cost a proportion of the indirect charges

of the Factory, such as Power, Unproductive Labour, Foreman's Wages, Works Managers' Salaries, etc., while the TOTAL COST is the Factory Cost plus the proportion of the Indirect Charges of Management generally, such as Office Expenses, etc. The indirect charges may be allocated to units of production in various ways, but frequently it will be found to be in the form of a percentage addition to the Cost. When this is the case the three Curves should move in complete agreement, but if some other basis is used the curves may be found to converge, and thus point to the fact that the Oncost is not being correctly allocated, or at least inquiry would need to be made to ascertain that the proportion added for these items is correct. A further chart which should be of great utility would show the ACTUAL Overhead expenditure in the works (*i.e.*, expenditure other than that in respect of Material and Direct Labour), and the amount allocated to the contracts or output. These curves should practically agree. If the curve showing the allocation of the expenditure, on whatever base it may be made, falls below that showing the actual expenditure for the same period, the addition to the Prime Cost is insufficient and the Factory Cost used for further work is incorrect, since the actual expenditure will have been more than that added on to the cost of the work, with the result that the selling price will not show the gross profit anticipated. On the other hand, if the curve for allocations is above that of expenditure the additions made are too great, and consequently the Factory Cost is too high, with the ultimate result that any calculations based upon it will be incorrect, and the business handicapped when quoting a price against its competitors. A similar plan would be adopted with regard to the General

Expenses and the allocation of the Oncost in arriving at the Total Cost ; and the same considerations would apply to this case as to the former. When the business is working to contracts these results can in a large measure be checked, if considered necessary, by a graph showing the accumulated output for a given period with the actual total cost thereof, and also the contract prices obtained. Fluctuations in wages can be kept in view by a chart showing the number of workpeople employed, and the total wages paid, while output can be compared with the number of employees in the factory or works, or, better still, with the Man-Hours worked. This should enable steps to be taken to maintain a steady output.

Other Cost diagrams which might be used are those showing the following information :—

(*a*) Cost of Manufacture compared with past experience, or a Standard of Cost based on such past experience.

(*b*) Cost of Raw Material in comparison with total Production Costs.

(*c*) Comparison of Direct and Indirect Expenses.

(*d*) Allocation and recoveries of Overhead Charges.

§ 12.—Transport Records.

Where a firm is engaged in Road Transport on any scale, valuable information can be obtained by recording the number of miles each van or wagon travels, the number of hours employed, the weight of the load carried, etc., and this information should also be analysed to show the proportion that the miles travelled when empty bear to the total mileage covered. By such means steps can be taken to ensure that all

similar vans or wagons carry similar loads and do not proceed half empty, or the rounds can be rearranged so that the numbers of miles run while empty can be considerably reduced. Similarly, by recording the costs of running, repairs, etc., the most economical vehicle can be found, and the personal skill of each driver indicated.

The following records might also be shown in graphic form :—

(*a*) The number of Wagon-Days operated.
(*b*) The number of Wagon-Days idle through Wagons being under repair.
(*c*) The number of Wagon-Days idle through lack of work.
(*d*) The number of Wagon-Hours operated per day.
(*e*) The number of Wagon-Hours during which loading or unloading was in progress.
(*f*) Miles run per gallon of petrol.
(*g*) Miles run per gallon of oil.
(*h*) Average number of miles per trip.

The graphic method is adaptable to any type of business or data, and consequently can be applied to a variety of commercial problems. A careful study of the object for which the data have been collected will enable the charts to be constructed to display this result with a clearness and ease unapproachable by any other method. The time taken is not greater, is indeed frequently much less, than that necessary to prepare tabulated statements, and, once the form of chart has been decided upon, it calls for no particular skill either in plotting or reading.

SYNOPSIS TO CHAPTER VIII.

METHODS OF COMPARISON.

§ 1.—Dispersion.

(a) Measuring Dispersion.

i. The Range.

ii. Absolute and Relative Dispersion.

iii. The Moments of Dispersion.

(1) The First Moment.

(2) The Second Moment.

(3) The Third Moment.

(4) The Modulus.

(5) The Quartile Measure and Co-efficient.

(6) The Lorenz Curve.

2.—Skewness.

(a) The First Measure and Co-efficient of Skewness.

(b) The Second Measure and Co-efficient of Skewness.

(c) The Third Measure and Co-efficient of Skewness.

CHAPTER VIII.

METHODS OF COMPARISON.

It has already been laid down in Chapter I that the ultimate end of Statistical Research is to enable comparison to be made, and various rules have already been enunciated with a view to enabling this to be carried out in a proper manner. It is now necessary to consider more advanced methods of comparing various groups of data, as well as measuring the relationship (if any) which exists between two or more different factors (or Variables).

§ 1.—Dispersion.

This term is used to indicate that within a given group of data there is no uniformity of size, or, in other words, that the items differ from one another in magnitude. If reference be made to the Table of Profits shown on page 77 it will be seen that no two items are of the same size, and that they vary from £35,000 in year 1, to £194,000 in the 20th year. If the variation from the actual size of a representative item (*e.g.*, the Arithmetic Average, the Median or the Mode) is small the dispersion is said to be slight, but when the sizes vary greatly from such standard item then the dispersion is said to be great. Suppose the profits

of one hundred businesses in the same trade and with the same Capital be taken, and the Profits vary from £10,000 to £10,100 then there is relative uniformity in their results, and the dispersion will obviously be but slight. If, however, the Profits varied from £1,000 to £15,000 there is very considerable dispersion.

(a) Measuring Dispersion.

There are various ways in which the dispersion may be measured with a view to studying the consistency of the results attained. It is, however, first necessary to consider what value is to be placed upon this consistency. For the sake of example let us take the profits earned by a Company. If in several consecutive Trading Periods the same amount of profit is made (say £25,000), then it will be agreed that there is absolute consistency in the results achieved, a state of affairs but seldom met with in commercial results. If, however, the profits varied between, say, £23,500 and £26,500, the results may be said to be fairly consistent. If we desire therefore to study consistency it is necessary to tabulate carefully and examine any deviations which occur in the data, since the wider these deviations are the less is the consistency of the results.

I. THE RANGE.

The simplest method of measuring the dispersion is to take the difference between the magnitude or characteristics of the extreme items of the data under review. In the Table of Profits set out on page 77 we find the Range of Dispersion to be £194,000—£35,000 = £159,000. If, however, an additional year's profits were included, and these amounted to £250,000 the Range of Dispersion would have been

increased by no less than £56,000, *i.e.*, from £159,000 to £215,000, but the average profits have only been increased from £67,750 to £76,428·57, an increase of only £8,678·57, a relatively small amount. It is obvious, therefore, that any unit of measurement which is capable of being affected to such a degree by the inclusion of additional items at the extremes is not satisfactory for use when a reasonable degree of accuracy is required. Again, the Range in two different groups of data may be the same, but the actual consistency of the results be quite different. Let us revert to the case already quoted, where the profits of a Company varied between £23,500 and £26,500, the Range being £3,000. Supposing another Company operating over the same period makes profits varying from £48,500 to £51,500. Here the Range of Dispersion is also £3,000, but the second Company is more consistently nearer its average profits than the first, for it had twice the chance of missing its average as compared with the first Company. It might have obtained results varying between £0 and £48,500, while the former only had a possible variation of £0 to £23,500. If, however, the second Company had shown results varying from £47,000 to £53,000 the consistency would have been the same as that of the first Company. This is clearly seen if we assume the following figures for three consecutive years :—

	1st Year.	2nd Year.	3rd Year.	Average.
Company A ..	£23,500	£25,000	£26,500	£25,000
Company B ..	£47,000	£50,000	£53,000	£50,000

Company A aims at achieving Profits amounting to £25,000 whereas the aim of Company B is to show Profits amounting to £50,000, or twice the amount

of the former. A variation of £1,500 in the case of A on an average of £25,000 is the same ratio of dispersion as that of a variation of £3,000 on an average of £50,000.

II ABSOLUTE AND RELATIVE DISPERSION.

It follows from the last example given that the absolute dispersion is not always to be relied upon. When measuring the absolute dispersion we ignore any question of the average, and may thus get a fallacious idea of the position. By using the relative dispersion results are frequently more reliable than those obtained by using the first named, but in order to measure this relative dispersion, with a view to comparing the deviations present in two or more groups, it is necessary to obtain a co-efficient of dispersion for each of the groups under review. The co-efficient represents the fraction of variation generally occurring in any group of data. *It is obtained by dividing the absolute measure of dispersion used by that magnitude which has been selected as representative of the data under review and from which the deviations have been measured. As no representative item is used in measuring the Range it follows that there can be no relative method of measurement in this case.*

III. THE MOMENTS OF DISPERSION.

Reference has been made to the method of measuring the deviation from the actual size of a representative item, *e.g.*, the Average, the Median or the Mode. In order that such deviations (or variations as they are sometimes known) may be accurately measured we use one of the **Moments of Dispersion** of which there are three—

(1) *The First Moment.* This is known as the average or mean deviation, *i.e.*, the sum of the deviations

from the type taken divided by the number of items under review, and in this case we ignore the signs which would normally appear before the deviations. The type selected may be either the Average, the Median or the Mode. Let d = the deviation from the average; d_M = the deviation from the Median; and d_X = the deviation from the Mode; and n = the number of examples; then:—

$\frac{\Sigma d}{n}$ = the first Moment of Dispersion from the Average;

$\frac{\Sigma d_M}{n}$ = the first Moment of Dispersion from the Median;

$\frac{\Sigma d_X}{n}$ = the first Moment of Dispersion from the Mode.

The application of the second formula to the following Frequency Table in which the data is composed of a Discrete Series, will illustrate the use of this Moment. The second formula is taken, *i.e.*, the Median is used as the Standard type.

(a).	(b).	(c).	(d)
Size of Item	Frequency.	Deviation from the Median	Products. (i.e., Deviation × Frequency)
9	3	5	15
10	6	4	24
11	8	3	24
12	11	2	22
13	16	1	16
14	18	0	0
15	15	1	15
16	9	2	18
17	7	3	21
	$n = 93$		$\Sigma d = 155$

From columns (*a*) and (*b*) of this Table it will be seen that the number of examples is 93, and that the

Median is therefore the 47th example. As this is a Discrete Series the size of the Median is 14. If, however, the data had comprised a Continuous Series, the magnitude of the Median would have been located by Interpolation as already explained in Chap. IV, § 4 (*b*). Having located the magnitude of the Median we record in Column (*c*) the deviation of the magnitude of each class from the selected type, and then multiply the figure so obtained by the number of items in each of the classes, this information being recorded in Column (*d*). Using the figures appearing as Totals of Columns (*b*) and (*d*) we get the following result:—

$\frac{\Sigma\, d_M}{n} = \frac{155}{93} = 1{\cdot}\dot{6}$. The sign δ is generally used to indicate this absolute measure of dispersion.

This calculation gives us the absolute measure of dispersion, and consequently the first of the factors required for the purpose of obtaining the co-efficient of dispersion referred to above. The other factor required being the quantity selected as representing a typically sized item. To calculate the co-efficient of dispersion we proceed as follows:—

Let δ = the absolute measure of deviation from the Average;

δ_M = the absolute measure of deviation from the Median;

δ_X = the absolute measure of deviation from the Mode;

and let a = the Average; M = the Median; and X = the Mode.

then $\frac{\delta}{a}$ = the co-efficient of dispersion from the Average;

$\frac{\delta_M}{M}$ = the co-efficient of dispersion from the Median.

and $\frac{\delta_X}{X}$ = that from the Mode.

Applying this to the above example we get—

$$\frac{\delta_M}{M} = \frac{1\cdot\dot{6}}{14} = \underline{\underline{0\cdot119.}}$$

The co-efficient of dispersion would be calculated for each group of data under review, and the results could then be easily compared. Its utility will be clearly seen if we take the results of the Companies, whose profits in three consecutive years are as shown in col. (*a*), calculating co-efficients of deviation for each of them.

(*a*).	(*b*).	(*c*).
Year.	Size of Item.	Deviation from the Average.
		COMPANY A.
1	£23,500	1,500
2	£25,000	0
3	£26,500	1,500
	Average £25,000	$\Sigma\, d = 3{,}000$
		COMPANY B.
1	£48,500	1,500
2	£50,000	0
3	£51,500	1,500
	Average £50,000	$\Sigma\, d = 3{,}000$

In each case there are three examples, hence in the case of A the absolute Measure of Dispersion is $\frac{3{,}000}{3} = 1{,}000$, and the co-efficient of dispersion is $\frac{1{,}000}{25{,}000} = \cdot04.$

In the case of B the absolute Measure of Dispersion is $\frac{3,000}{3} = 1,000$ the same as in A, but the co-efficient in this case is $\frac{1,000}{50,000} = \cdot 02$. It is at once apparent that the co-efficient of dispersion corrects the wrong impression which the absolute measure of dispersion gives, and therefore enables the relative consistency in the results to be studied more accurately. Obviously the lower the co-efficient of dispersion is, the more regular is the consistency in the results achieved.

(2) *The Second Moment.* The absolute measure of dispersion from the average in this case is known as the **Standard Deviation,** and is the one most frequently used. The moment is the sum of the deviations when squared, divided by the number of the items, *viz.*, $\frac{\Sigma d^2}{n}$. The Standard Deviation is the square root of this second moment, *i.e.*, $\sqrt{\frac{\Sigma d^2}{n}} = \sigma$, and its working is shown from the following Frequency Table of a Discrete Series.

(*a*).	(*b*).	(*c*).	(*d*).	(*e*).	(*f*).
Size of Item.	Frequency.	Sum of Sizes. [Col. (*a*) × (*b*)].	Deviation from Average.	d^2	Products. [Col. (*b*) × (*e*)].
6	3	18	− 3	9	27
7	6	42	− 2	4	24
8	9	72	− 1	1	9
9	13	117	0	0	0
10	8	80	+ 1	1	8
11	5	55	+ 2	4	20
12	4	48	+ 3	9	36
	$n = 48$	432			$\Sigma d^2 = 124$

The average to be used is the Weighted Arithmetic Average and is therefore $\frac{432}{48} = 9$. The Total of Column (f) gives us the sum of the deviations when squared, and from the average and this total the Standard Deviation can be obtained, *e.g.*,

$$\sqrt{\frac{\Sigma d^2}{n}} = \sqrt{\frac{124}{48}} = \sqrt{2{\cdot}58\dot{3}} = 1{\cdot}607.$$

From this Standard Deviation the co-efficient of dispersion is calculated in the same manner as before,

viz., $\frac{\sigma}{a} = \frac{1{\cdot}607}{9} = 0{\cdot}178\dot{5}$

Short-Cut Method. The above method of arriving at the Standard Deviation is the simplest only when the average of the items is a whole number, or when the deviations from such average do not contain fractions. It frequently happens, however, that the average or the deviation therefrom is a fraction, or contains a fraction, and in such cases the work of calculating the Deviations, and squaring them entails considerable labour, thus leaving loopholes for mistakes. In such cases the following Short-Cut Method will be found very useful, as well as result in a considerable saving of time. It is based on the mathematical proposition that *the sum of the squares of the deviations from the Arithmetic Average is a minimum*, and this statement must be accepted, as it is no part of the scope of this manual to demonstrate the truth of generally accepted mathematical formulæ. The method is as follows :— **Calculate the average size of the items under review ; select a whole number approximating this true average ; compute the deviations from such assumed average ; square each of these deviations and summate the results. Subtract**

from such summation n times the square of the difference between the assumed average and the true average; divide by n; extract the square root of the result. This may be expressed in the following formulæ; when x = the assumed average and a = *actual average, viz*:—

$$\sigma = \sqrt{\frac{\Sigma d_x^2 - n\,(a-x)^2}{n}}$$

The following example will illustrate the method of procedure:—

(a).	(b).	(c).	(d).	(e).	(f).
Size of Item.	Frequency.	Sum of Sizes. [Col. (a) × (b)]	Deviations from Assumed Average. d_x	d_x^2	Products. [Col (e) × (b)]
4	6	24	− 4	16	96
5	8	40	− 3	9	72
6	10	60	− 2	4	40
7	15	105	− 1	1	15
8	14	112	0	0	0
9	12	108	+ 1	1	12
10	9	90	+ 2	4	36
	$n = 74$	539			$\Sigma d_x^2 = 271$

The true average $= \dfrac{539}{74} = 7{\cdot}284 = a.$

Let the assumed average $= 8 = x$

then $(a - x)^2 = (7{\cdot}284 - 8)^2 = (-{\cdot}716)^2 = {\cdot}5126$

then $n\ (a - x)^2 = 74 \times {\cdot}5126 = 37{\cdot}93.$

$$\therefore \sqrt{\frac{\Sigma d_x^2 - n(a-x)^2}{n}} = \sqrt{\frac{271 - 37{\cdot}93}{74}} = \sqrt{\frac{233{\cdot}07}{74}}$$

$$= \sqrt{3{\cdot}151} = \underline{1{\cdot}775}$$

If we take an assumed average of 7, the result will be the same, In this case, columns (d), (e) and (f) will appear as follows:—

(d)	(e)	(f)
− 3	9	54
− 2	4	32
− 1	1	10
0	0	0
+ 1	1	14
+ 2	4	48
+ 3	9	81
		$\Sigma d_x^2 = 239$

then $(a - x)_x^2 = (7 \cdot 284 - 7)^2 = (\cdot 284)^2 = \cdot 08 +$, then $n(a - x)^2 = 74 \times \cdot 08 = 5 \cdot 93$ approx.

$$\therefore \sqrt{\frac{\Sigma d_x^2 - n\,(a-x)^2}{n}} = \sqrt{\frac{239 - 5 \cdot 93}{74}} = \sqrt{\frac{233 \cdot 07}{74}}$$

which gives the same result as before.

The Standard co-efficient of dispersion is found in the same way as before, and in the above case would be

$$\frac{\sigma}{a} = \frac{1 \cdot 775}{7 \cdot 284} = 0 \cdot 2449$$

The Standard Deviation and the resulting co-efficient of dispersion have been but little used in Economic Research, but is very largely used by biologists. From a Commercial point of view the Average Deviation is in many ways preferable, for the squaring of the variations gives more weight to extreme items, a measure which is but seldom required, seeing that it is the results of the modal class which will particularly interest Commercial Statisticians. Moreover, there is considerable work involved in computing the Standard Deviation, and this alone makes it less desirable for commercial work where rapidity in the preparation of the results from the data obtained is a very important factor.

(3) *The Third Moment.* The formula for this measure of dispersion is the sum of the cubes of the deviations divided by the number of items, *i.e.*, $\frac{\Sigma d^3}{n}$.

The factors which tend to make the Standard Deviation unsuitable for commercial work would apply in even a greater degree to this measure of dispersion, and hence it need not be further considered.

(4) *The Modulus.* This is a measure of dispersion, used in advanced Statistical work, and, like the Standard Deviation, is based on the Second Moment. It is usually represented by c, and the formula for its calculation is—

$$c = \sqrt{\frac{2 \Sigma d^2}{n}}$$

Here again there is no need for additional consideration, as this measure of dispersion will be but seldom met with, or required, in commercial research.

(5) *The Quartile Measure and Co-efficient of Dispersion.* This is based upon the deviations between the Upper and Lower Quartiles, and ignores the individual deviations of the items in the data from the chosen standard. It gives a general idea of the dispersion without requiring a great amount of detailed work such as the other measures of dispersion need. It will be immediately recognised that one half of the data falls between the two quartiles, and, since in most statistical data extreme magnitudes tend to be comparatively rare, then such examples as lie between the Upper and Lower Quartiles should be more or less representative of the whole. If we double the number of items in an array, and such additional items are more or less evenly spread over it, then the distance between the Upper and the Lower Quartiles will also be approximately doubled.

The Quartile Deviation is easily calculated from the following formula, in which Q_l = the magnitude

of the Lower Quartile; and Q_u = the magnitude of the Upper Quartile, *viz.*—

$$\frac{Q_u - Q_l}{2}$$

To obtain the Quartile co-efficient of dispersion it is necessary, as before, to divide the Quartile measure of dispersion by a number representing the size of a typical item under review. It is usually considered that the mean of the magnitudes of the Upper and Lower Quartiles is the most satisfactory denominator, *viz.*—

$$\frac{Q_u + Q_l}{2}$$

The co-efficient of dispersion then =

$$\frac{\frac{Q_u - Q_l}{2}}{\frac{Q_u + Q_l}{2}} = \frac{Q_u - Q_l}{Q_u + Q_l}$$

Example. If we turn to the Frequency Table appearing on page 224, we shall find that the Lower Quartile is of the size 12, while the Upper Quartile measures 15. Using the magnitudes so obtained the Quartile Measure of Dispersion =

$$\frac{15 - 12}{2} = \frac{3}{2} = \underline{\underline{1{\cdot}5}}.$$

The Quartile co-efficient of Dispersion therefore =

$$\frac{\frac{15 - 12}{2}}{\frac{15 + 12}{2}} = \frac{\frac{3}{2}}{\frac{27}{2}} = \frac{1{\cdot}5}{13{\cdot}5}$$

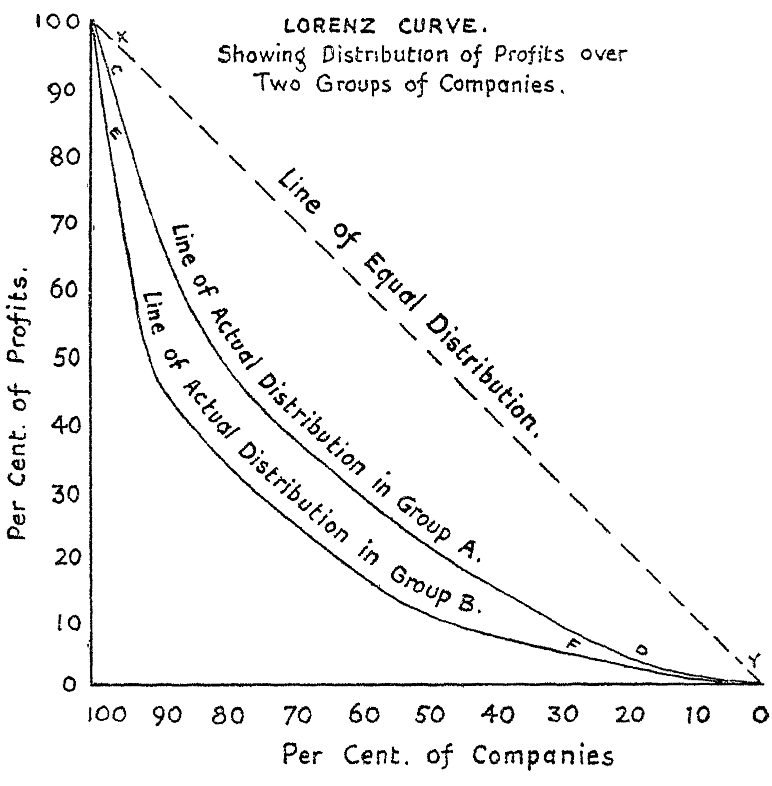
$$= \underline{\underline{0{\cdot}\dot{1}}}$$

(6) *The Lorenz Curve.* The graphic method can also be used for the measuring of divergencies from the average. The best method is that adopted by Dr. Lorenz for the measuring of the distribution of wealth, the curve showing the distribution being known as the **Lorenz Curve.** It takes the form of a cumulative percentage curve, combining the percentage of items under review with the percentage of wealth or other factor distributed among such items. This curve should be very useful in commercial statistical work for comparing the distribution of profits over different groups of businesses. The following example illustrates the method of construction of the curves shown in the accompanying Diagram. Two groups of businesses, A and B, are to be compared with a view to studying how the profits earned in each group have deviated from the average, and from one another. *

Total Amount of Profits earned by Companies in each Division.	No. of Companies in each Division.	
	Group A.	Group B.
£		
600	6	1
2,500	11	19
6,000	13	26
8,400	14	14
10,500	15	14
15,000	17	13
17,000	10	6
40,000	14	7
£100,000	100	100

It is first necessary to turn these results into a Cumulative Percentage Table as below. It should be

* For the sake of convenience the number of Companies in each of the groups has been made 100, and the profits earned £100,000. If the numbers and amounts did not amount to 100, or a multiple thereof, percentages would have to be calculated.


LORENZ CURVE.
Showing Distribution of Profits over Two Groups of Companies.
Line of Equal Distribution.
Line of Actual Distribution in Group A.
Line of Actual Distribution in Group B.
X
C
E
F
D
Y
100
90
80
70
60
50
40
30
20
10
0
Per Cent. of Profits.
Per Cent. of Companies

noted that in the example the cumulative frequencies also give the percentages, otherwise additional columns would be provided to show these.

Profits			Group A.		Group B	
Amount of Profits.	Cumulative Profits.	Percentage of Profits to Total.	No. of Companies.	Cumulative No of Cos and % of Total.	No of Companies	Cumulative No. of Cos and % of Total.
(a)	(b).	(c).	(d)	(e)	(f).	(g)
£	£					
600	600	0 6	6	6	1	1
2,500	3,100	3 1	11	17	19	20
6,000	9,100	9·1	13	30	26	46
8,400	17,500	17·5	14	44	14	60
10,500	28,000	28·0	15	59	14	74
15,000	43,000	43·0	17	76	13	87
17,000	60,000	60 0	10	86	6	93
40,000	100,000	100·0	14	100	7	100

In plotting the percentages shown in Column (*c*) the vertical scale is used. As will be seen this reads upwards from 0 to 100 per cent. The data in Columns (*e*) and (*g*) are plotted from the horizontal scale which reads from right to left. The curve CD represents the distribution of Profits in Group A, while that marked EF shows the distribution in Group B. If the distribution of profits over either of the groups was even, *i.e.*, if all of the examples made the same amount of profit, the straight line XY would represent the distribution, therefore the divergence between this line of equal distribution and the curve of actual distribution will picture the deviations from the average which are present in the data under review. Similarly, a comparison of the curves representing the actual distribution in the different groups under investigation will show the difference which exists between the two groups. This method does not furnish us with any numerical result which may be used for measuring the dispersion, and should therefore be used in conjunction

with one of the co-efficients of dispersion whenever a detailed study of the deviations is required.

There is obviously a tendency for the curves to coincide at the extremities, and it is therefore usual to plot the two extremes on separate charts, using a much larger scale in order that the deviations of the extreme items may be properly analysed and studied.

§ 2.—Skewness.

As already explained in Chapter VI, Skewness is present when the dispersion of items in a group is not symmetrical, and results in the average and the Median being pulled away from the Mode, near which they are both located when the deviation is of an even character. It frequently becomes necessary, however, to study the Skewness present in two groups of varying data of the same nature, and, in order to do this, the Skewness must in each case be converted into a numerical quantity, for the curves themselves will usually be such that they cannot be compared. Co-efficients of Skewness must also be calculated and used, for the same deficiencies which are present in the Measures of Dispersion are also present in those of Skewness. In the case of Dispersion we desire to ascertain the deviation from the chosen standard which exists in the data being investigated, whereas in the case of Skewness it is necessary to measure the amount by which the deviations on one side of the standard taken exceed those on the other. In obtaining the co-efficient of Skewness, therefore, the divisor used must always be some measure of deviation, and not the chosen standard itself as is the case in the co-efficient of dispersion.

(a) The First Measure and Co-efficient of Skewness.

The simplest measure in this case is the distance the Arithmetic Average is diverted from the Mode, and is

found by subtracting the magnitude of the Mode from the Arithmetic Average, *i.e.*, $a - x$, when a = the Arithmetic Average, and x = the magnitude of the Mode. The symbol j is used to indicate the co-efficient of Skewness, and the co-efficient based on the above measure will be—

$$j = \frac{a - x}{\delta_x}$$

Very little difference will result if the average deviation from the Arithmetic Average be used in place of the average deviation from the Mode used in the formula given above, provided of course that the same divisor is used in all cases which it is desired to compare.*

As we have already seen in Chapter IV, it frequently happens that the Mode is not very clearly defined, and hence cannot be used with any degree of certainty. In such cases the best measure which can be used is $a - M$ (*i.e.*, the difference between the Arithmetic Average and the Median), and the co-efficient of Skewness would then be calculated from $\frac{a - M}{\delta_M}$. It must however be remembered that the Median will have been moved at the same time, though to a lesser degree, than the Arithmetic Average, and it therefore follows that this Measure and co-efficient of Skewness will not be so satisfactory in use as the one based on the Mode, when this can be clearly identified, for this is seldom affected by additional examples.

(b) The Second Measure and Co-efficient of Skewness.

When Skewness occurs in an array, the magnitude of the Quartile nearer to the extended end of the

* As Skewness may occur on either side of the Mode it is obvious that the average may be smaller in magnitude than the Mode, and hence the result of the above computation would show a minus answer. This minus result would indicate Skewness below the Mode, whereas in the other case the Skewness would occur between the Mode and the upper extremity.

curve is increased in the direction of the magnitude of that extremity to a greater degree than is the magnitude of the other Quartile. The result of this is that the magnitude of the Median will no longer be approximately mid-way between the magnitude of the Upper and Lower Quartiles. In the data tabulated in Table P (page 146) the magnitude of the Median will be found to be £2,165; that of the Lower Quartile is £1,547·14; while the Upper Quartile Measures £2,795. The difference in the magnitudes of the Median and Lower Quartile is therefore £2,165·14 – £1,547 = £618·14; while between the Upper Quartile and the Median there is a difference of £2,795 – £2,165·14 = £629·86. If we now turn to Table Q (page 150), from which the Skew Curve in Diagram 3a was plotted, we find that the magnitude of the Median is £2,547·06; the Lower Quartile measures £1,765·8; and the Upper Quartile is £3,477·7. In this case, therefore, the difference between the magnitudes of Median and the Lower Quartile is £781·26, but between the Upper Quartile and the Median the difference measures no less than £930·64, a greatly increased difference as a result of the Skewness. The second measure of Skewness is based on the fact that the Median moves but little as compared with the Quartile nearest to the extremity on which the Skewness lies. *It is calculated by adding the magnitudes of the Upper and Lower Quartiles, and subtracting from the summation twice the magnitude of the Median, i.e.,* $Q_u + Q_l - 2M$. To obtain the co-efficient in this case we divide the above measure by the Quartile Measure of Deviation, which gives us the formula:—

$$j = \frac{Q_u + Q_l - 2M}{\frac{Q_u - Q_l}{2}}$$

Applying this to Table Q already quoted—

$$j = \frac{3477{\cdot}7 + 1765{\cdot}8 - 2\,(2547{\cdot}06)}{\dfrac{3477{\cdot}7 - 1765{\cdot}8}{2}}$$

$$= \frac{149{\cdot}38}{855{\cdot}85}$$

$$= \underline{0{\cdot}1747.}$$

If the co-efficient is preceded by a plus sign the Skewness occurs between the Median and the upper extremity, while if a minus sign precedes the co-efficient then the Skewness is in the lower half of the array. This co-efficient of Skewness enjoys the same advantages and suffers from the same disadvantages as the Quartile co-efficient of dispersion. The principal objection is the fact that the extreme items are not considered at all, but, as already pointed out, for commercial use a study of the examples located at the position of greatest density is frequently sufficient.

(c) The Third Measure and Co-efficient of Skewness.

The formula used for measuring the Skewness in this case is—

$$\sqrt[3]{\frac{\Sigma\, d^3}{n}}$$

The solution of this requires considerable calculation, and moreover gives increased weight to the magnitudes of the extreme items, and hence is unsuitable in every way for commercial work. To obtain the co-efficient of Skewness either of the measures of deviation may be used, but the Standard Deviation is that generally adopted. In such a case the formula for the calculation of the co-efficient would be—

$$j = \frac{\sqrt[3]{\dfrac{\Sigma\, d^3}{n}}}{\sigma}$$

SYNOPSIS TO CHAPTER IX.

INDEX NUMBERS IN GENERAL.

§ 1.—The Nature and Purpose of Index Numbers.

2.—The Limitations of Index Numbers.

3.—The Relative.

4.—The Arithmetic Mean (A.M.).

5.—The Geometric Mean (G.M.).

6.—The A.M. and G.M. Compared.

7.—The Chain Base Method.

8.—The Weighting System.

9.—The Aggregative Method.

10.—Changing the Base Year.

11.—Index Numbers of Quantity.

CHAPTER IX.

INDEX NUMBERS IN GENERAL.

§ 1.—The Nature and Purpose of Index Numbers.

If a business man is faced with a table of prices or production figures, he may ask " What is the general trend of these figures ? How much have prices risen, on the average, in the last year, or in the last ten years ? Is production, taken as a whole, increasing satisfactorily ? Is output per man-hour increasing, and if so, by how much ? " Politicians have similar problems to consider for the nation as a whole, together with such matters as "terms of trade." The man in the street would also like to know how much his cost of living has risen in recent years and whether his wages are keeping pace with rising prices. Whether they should do so is another matter with which this book is not concerned.

These and many other questions can be answered by means of *index numbers*, or more briefly, indices. An index number is, as its name suggests, an indicator of the general trend of a set of figures, and it is usually expressed as a percentage. Thus, if it is estimated that prices of a certain commodity-group have risen, on the average, 70% since 1938, the price index is said to be 170, based on 1938 as 100.

An index number must be compiled with reference to some standard period called the base period. It is usually convenient to take a year, known as the base year, and in the example mentioned above the base year is 1938. This particular year is often chosen for

this purpose, being the last pre-war year and in most respects fairly normal. It is possible to take a month, a day, or any other period as base, but the principle is the same. If it is difficult to select a typical year, the average of a series of years may be taken, as in *The Statist* Index Numbers (*see* Chap. X, §2).

The value of the index number in the base period is called the base index, and is nearly always 100, but there are exceptions. The South African Index Number of Retail Prices, for instance, has a base of 1,000, which has the advantage of giving three or four significant figures without requiring decimals. In the early days of *The Economist* Wholesale Price Index, the base was 2,200 (*see* Chap. X, § 3), but that eminent journal has long ago mended its ways.

§ 2.—The Limitations of Index Numbers.

Convenient as index numbers are, they have several limitations. In the first place, an index is generally based on a sample. For example, one cannot include all commodities in common use in an index of retail prices, and the principle of sampling must therefore be adopted. Of course, the more items that can be introduced the more reliable will be the index number, as it will be less affected by any abnormal fluctuations in the price of one or more items, but one may be limited by the difficulty or cost of collecting the data, or the necessity of compiling and publishing the index as quickly as possible.

Secondly, an index of prices or production may not take into account variations in quality, which may be considerable. Naturally an article of superior quality will cost more at any given time than an inferior article, and a rise in the price index may be due to an

improvement in quality, not to a rise in prices, but very often there is no information on this point. Thus the " Annual Abstract of Statistics " for 1935-46 gives average prices for commodities such as cheese and wool, regardless of grade or quality. Again, during the 1939-45 War there were much smaller sales of expensive paints than before the war, consequently any price index that merely took an average price for all grades of paint would understate the general increase in prices.

A somewhat similar point arises in connection with cost of living index numbers. How can one compare the cost of living in different countries or different periods when habits of life vary so much ? How can the standard of living of a man who drinks and smokes heavily be compared with that of a man who does neither, or that of an Englishman with that of an Eskimo ? There is no perfect solution to these problems. Furthermore, one cannot legitimately compare prices in, say, 1948 with those of 1938 without taking into account the effect of subsidies, rationing and shortages, and remembering that restrictions on consumer's choice are an economic loss.

Finally, an index number can be calculated in many different ways, and different methods give different answers. On the whole, however, unless conditions are changing violently, the results of the various methods will agree fairly closely and should at least show the same general trend. Very often that is all that really matters.

§ 3.—The Relative.

In the following sections on the theory of index numbers and methods of construction it will be con-

venient to refer to a price index, with a base period of a year unless otherwise stated. The methods apply equally, for the most part, to indices of wages, profits, production, etc., although some of these have their own peculiar problems, which will be considered later on.

It is usual to begin by expressing the price of each commodity as a percentage of its average price for the base year. This percentage is sometimes called a *relative*, since it expresses the current price relatively to the base price. For example, Table 304 of the "Annual Abstract of Statistics" for 1935-46 gives the annual average price of plums in England and Wales as follows :—

Year	Price per cwt. s.	d.	Year	Price per cwt. s	d.
1935	37	6	1941	57	6
1936	14	3	1942	31	6
1937	19	0	1943	34	6
1938	35	0	1944	34	4
1939 .. .	14	6	1945	38	1
1940	20	6	1946	37	8

Table 305 shows these prices as indices, or relatives. On account of the somewhat erratic variations from year to year, the average for the three years 1936 to 1938 is taken as the base, in this case 22s. 9d., and each price is expressed as a percentage of this amount. The result, which the student should verify for himself, is as follows :—

Price relatives (1936/38 = 100).

Year	Relative	Year	Relative	Year	Relative
1935 ..	165	1939 ..	63	1943 ..	151
1936 ..	63	1940 ..	90	1944 ..	150
1937 ..	84	1941 ..	253	1945 ..	168
1938 ..	153	1942 ..	138	1946 ..	165

It will be noticed that the three relatives for 1936 to 1938 total 300 and average 100, as might be expected. For some commodities however, the total is 299 or 301, owing to rounding off.

§ 4.—The Arithmetic Mean (A.M.).

The next step is to combine a number of price relatives into a single price index. The simplest method is to take the arithmetic mean (A.M.) of the relatives for any particular year.

Consider three commodities A, B and C, and suppose their relatives for 1947 based on 1940 as 100 are 105, 112 and 103 respectively. Taking the simple A.M., the price index for 1947 is

$$\tfrac{1}{3}\,(105 + 112 + 103) = 106{\cdot}7$$

correct to one place of decimals, which is usually sufficient.

If, however, A, B and C are not regarded as equally important, weights can be chosen corresponding to their relative importance. For instance, suitable weights might be numbers roughly proportional to the values of the sales of A, B and C in the base year. If these weights are 2, 3 and 5, the price index becomes

$$\frac{2 \times 105 + 3 \times 112 + 5 \times 103}{2 + 3 + 5} = 106{\cdot}1$$

In general, if P_1, P_2 . . . P_n are average prices for the base year and P_1', P_2' .... P_n' are average prices for the current year, the relatives (omitting the factor 100) are

$$\frac{P_1'}{P_1},\ \frac{P_2'}{P_2},\ \ldots\ldots\ \frac{P_n'}{P_n}\ ;$$

the simple, or unweighted, A.M. is

$$\frac{1}{n}\left(\frac{P_1'}{P_1} + \frac{P_2'}{P_2} + \ldots\ldots + \frac{P_n'}{P_n}\right) \qquad \ldots. \qquad (1)$$

and the weighted A.M., using weights W_1, W_2, ... W_n, is

$$\frac{W_1\dfrac{P_1'}{P_1} + W_2\dfrac{P_2'}{P_2} + \ldots\ldots + W_n\dfrac{P_n'}{P_n}}{W_1 + W_2 + \ldots\ldots\ldots\ldots + W_n} \qquad \ldots.\ \ldots\ldots \qquad (2)$$

or, using the Σ (sigma) notation, in which Σx denotes "the sum of all quantities like x," the simple A.M. is

$$\frac{1}{n}\Sigma\left(\frac{P'}{P}\right) \quad \ldots \quad (3)$$

and the weighted A.M. is

$$\frac{\Sigma\left(W\frac{P'}{P}\right)}{\Sigma W} \quad \ldots \quad (4)$$

Example 1.

Construct a price index based on 1935 from the table shown below, using a weighted A.M. with weights 2, 4, 7, and 7 respectively.

Year	A	B	C	D
	s/cwt.	s/cwt.	s/cwt.	s/cwt.
1935	70	25	20	60
1936	67	16	14	45
1937	58	21	22	65
1938	90	28	17	75
1939	65	18	18	52

First obtain the relatives for A, B, C and D, taking 1935 as 100 in each case. Then multiply the figures by 2, 4, 7 and 7, respectively, for each year and divide the totals by 20, the sum of the weights. The result (which the student should verify) is shown below.

Year	A	B	C	D	Index
1935	100	100	100	100	100
1936	96	64	70	75	73
1937	83	84	110	108	101
1938	129	112	85	125	109
1939	93	72	90	87	86

§ 5.—The Geometric Mean (G.M.).

The geometric mean is often used in preference to the arithmetic mean. As will appear later, it has certain theoretical advantages, particularly in price indices, and is being increasingly used in this connection. Its chief disadvantage is the laborious

calculations it often involves, as it almost invariably requires logarithms.

Returning to the simple illustration of the previous section, the G.M. of three price relatives 105, 112 and 103 is

$$\sqrt[3]{105 \times 112 \times 103},$$

which is found to be 106·6 correct to one place of decimals, slightly less than the value found by the A.M.

If there are n items, of equal "weight," with relatives

$$\frac{P_1'}{P_1}, \frac{P_2'}{P_2}, \ldots . \frac{P_n'}{P_n}$$

as before, the index I obtained from the G.M., omitting the factor 100 for convenience, is given by

$$I = \sqrt[n]{\frac{P_1'}{P_1} \times \frac{P_2'}{P_2} \times \ldots \ldots \times \frac{P_n'}{P_n}} \quad \ldots \quad \ldots \quad (5)$$

Taking logarithms,

$$\log I = \frac{1}{n} \Sigma \log \left(\frac{P'}{P}\right) \quad \ldots\ldots \quad \ldots\ldots \quad (6)$$

With the G.M. it is not necessary, as it is with the A.M., to calculate each relative separately, as the actual prices for either year can be multiplied together. Formula (5) can be expressed

$$I = \sqrt[n]{\frac{P_1' \times P_2' \times \ldots \times P_n'}{P_1 \times P_2 \times \ldots \times P_n}} \quad \ldots\ldots\ldots \quad (5\text{A})$$

and formula (6) as

$$\log I = \frac{1}{n} (\Sigma \log P' - \Sigma \log P) \quad \ldots \quad \ldots\ldots\ldots \quad (6\text{A})$$

(*N.B.*—The student who finds he has forgotten how to use logarithms will be well advised to revise them.)

Example 2.

Using a simple G.M., obtain from the following table a price index for 1946 based on 1938.

Commodity	Units	1938 price	1946 price
A	s/ton	52	123
B	d/lb.	12·7	21·8
C	d/gall	36 9	74·1
D	£/flask	4 85	16 90

The simplest method of calculation, which the student should study carefully, is as follows :—

	1938		1946	
Commodity	P	$\log P$	P'	$\log P'$
A	52	1·7160	123	2·0899
B	12·7	1·1038	21·8	1·3385
C	36·9	1 5670	74·1	1·8698
D	4 85	0 6857	16 90	1·2279
Total		5·0725		6·5261
				5·0725
				4)1·4536
				0·3634

$$\therefore I = \text{anti-log } 0{\cdot}3634$$
$$= 2{\cdot}309$$

or, taking 1938 as 100, the index for 1946 is 230·9.

Note that it makes no difference what the units are, provided the price of any particular commodity is expressed in the same units each year. Thus, commodity B might have been quoted in £/cwt. or £/ton, but the price relative would be unaltered, as both prices would be multiplied by the same factor. This applies equally to the A.M. In the case of the G.M. both logarithm totals would be increased or diminished by the same amount, and their difference would be unchanged.

If the items are not of equal importance they can be weighted in much the same way as in the A.M., except that the weights are no longer factors, but indices in the

usual algebraic sense of the word. Suppose, for instance, there are two commodities A and B, which are to be assigned weights of 2 and 3. This is done by treating A and B as if they consisted of 2 and 3 items respectively, all equally weighted, so that the index is given by

$$I = \sqrt[5]{\frac{P_a'}{P_a} \times \frac{P_a'}{P_a} \times \frac{P_b'}{P_b} \times \frac{P_b'}{P_b} \times \frac{P_b'}{P_b}}$$

$$= \sqrt[5]{\left(\frac{P_a'}{P_a}\right)^2 \left(\frac{P_b'}{P_b}\right)^3}$$

In general, the weighted G.M. of n items with relatives $\frac{P_1'}{P_1}$, etc., and weights W_1, W_2, ... W_n is given by

$$I = \sqrt[(\Sigma W)]{\left(\frac{P_1'}{P_1}\right)^{W_1} \left(\frac{P_2'}{P_2}\right)^{W_2} \cdots \left(\frac{P_n'}{P_n}\right)^{W_n}} \quad .. \ (7)$$

Taking logarithms as before,

$$\log I = \frac{W_1 \log \frac{P_1'}{P_1} + W_2 \log \frac{P_2'}{P_2} + \ldots\ldots + W_n \log \frac{P_n'}{P_n}}{W_1 + W_2 + \ldots\ldots + W_n}$$

$$= \frac{\Sigma \left(W \log \frac{P'}{P}\right)}{\Sigma W} \quad \ldots\ldots (8)$$

The student will see, by comparing these formulae with those for the A.M., that the G.M. is in effect calculated by taking the A.M. of the logarithms of the relatives instead of the A.M. of the relatives themselves.

Example 3.

With the same data as in Example 2, construct a price index for 1946 based on 1938 by means of a weighted G.M., giving A, B, C, D weights 3, 2, 1 and 5, respectively.

The table is now a little more elaborate, and again the student should study it carefully. Here it is best to subtract $\log P$ from $\log P'$ for each commodity before multiplying by the appropriate weight.

Commodity	P	$\log P$	P'	$\log P'$	$\log \frac{P'}{P}$	W		$W \log \frac{P'}{P}$
A	52	1·7160	123	2·0899	0·3739 ×	3	=	1·1217
B	12 7	1·1038	21 8	1 3385	0·2347 ×	2	=	0·4694
C	36·9	1·5670	74 1	1 8698	0 3028 ×	1	=	0·3028
D	4·85	0·6857	16·90	1·2279	0·5422 ×	5	=	2 7110
Total								11)4 6049
								0·4186

$$\therefore I = \text{anti-log. } 0{\cdot}4186 = 2{\cdot}622,$$
or, as a percentage, 262·2.

§ 6.—The A.M. and G.M. Compared.

It is evident that the A.M. is much simpler to calculate and more readily understood than the G.M., but the latter, as has already been said, possesses certain advantages over the A.M., the chief of which can now be explained.

Suppose the problem in Example 2 had been to construct a price index, not for 1946 based on 1938, but for 1938 based on 1946. The working would be the same as far as the addition of the columns of logarithms. It would then proceed as follows :—

$$\begin{array}{lr} & 5{\cdot}0725 \\ & 6{\cdot}5261 \\ \hline I = \text{anti-log } \bar{1}{\cdot}6366 & 4\ )\bar{2}{\cdot}5464 \\ \hline \quad = 0{\cdot}4331 & \bar{1}\ 6366 \\ \hline \end{array}$$

Obviously, from the way they have been worked out, the two final logarithms 0·3634 (obtained in Example 2) and $\bar{1}{\cdot}6366$ add up to zero, so their anti-logs are reciprocals. In other words, the index for 1946 based

on 1938 is the reciprocal (allowing for the factor 100) of the index for 1938 based on 1946. In this case, allowing for rounding off, $2{\cdot}309 \times 0{\cdot}4331 = 1$.

The G.M., therefore, shows the same relative movement between the two years, whichever is taken as the base year. This property, known as reversibility, holds good equally for the weighted G.M. provided the weights are the same in both cases. It does not hold, however, for the A.M., unweighted or weighted, as can be shown by a simple example, purposely exaggerated.

Example 4.

One commodity, A, doubles its price, and another, B, halves its price between 1938 and 1947. Using a simple A.M., calculate a price index (i) for 1947 based on 1938 and (ii) for 1938 based on 1947.

(i) Taking 1938 as 100,
relative of A = 200
" B = 50
$\therefore I = \frac{1}{2}(200 + 50) = 125$

(ii) Taking 1947 as 100,
relative of A = 50
" B = 200
$\therefore I' = \frac{1}{2}(50 + 200) = 125.$

By the first method, therefore, the general price-level was 25% higher in 1947 than in 1938, whereas by the second method it was 25% higher in 1938 than in 1947! The explanation is that the A.M. is biased, in that it tends to give too high a result. If it was right to give A and B equal weight in 1938, it must be wrong to give them equal weight in 1947 when the price of A has risen so much compared with B and sales of A will probably be very much less. In other words, the higher relatives are over-weighted.

The G.M. has the effect of giving less weight to high relatives and so removing the bias that exists in the A.M. It can be shown by a little algebra that a G.M. is always less than the A.M. with the same weights, and that the product of the two indices obtained by the A.M., using each year in turn as base year, always exceeds unity unless, of course, all the relatives are equal, when the A.M. and G.M. are also equal.

§ 7.—The Chain Base Method.

This is a convenient point at which to introduce the Chain Base Method, although it might be more logical to leave it until the Aggregative Method has been dealt with.

So far the discussion has been confined to the Fixed Base Method, in which each year is compared directly with the base year. Sometimes, however, this is impracticable, and it is more convenient to use the Chain Base Method, in which each year is compared with the previous year and the resulting indices multiplied together successively to link up with the base year.

Thus if the index for 1946 based on 1945 is 107·5 and the index for 1947 based on 1946 is 104·8, the index for 1947 based on 1945 is

$$107{\cdot}5 \times \frac{104{\cdot}8}{100} = 112{\cdot}66,$$

and so on; or looking at it another way, if the price-level rose 7·5% in 1946 and another 4·8% (on 1946) in 1947, the total rise since 1945 is

$$7{\cdot}5\% + 4{\cdot}8\% \text{ of } 107{\cdot}5\% = 12{\cdot}66\%.$$

This method has the advantage that new items can be introduced in their second year and old items

retained in the index as long as possible. For example, suppose a firm manufacturing plastics, which began making a new synthetic resin in 1942, has a price index based on 1935. On the fixed base method, the new resin cannot be brought into the index at all, for there is no price for 1935 with which to compare current prices; but on the chain base method, it can be introduced in 1943. Similarly, a commodity which was originally important but ceased to be manufactured in, say, 1946, can be dropped from the index in that year without the distorting effect that might have been caused by comparing 1946 directly with 1935 and omitting products which, perhaps, constituted the bulk of the earlier year's sales. When the pattern of production or sales is changing rapidly, the chain base method is generally preferable to the fixed base method.

Even when the same items are retained throughout, the two methods will normally give slightly different results if the A.M. is used. With the G.M., however, they give the same results provided the same weights are used throughout. Let I_{12} denote the index (omitting the factor 100) for year 2 based on year 1, I_{23} the index for year 3 based on year 2, and I_{13} the index for year 3 based directly on year 1. Then the index for year 3 based on year 1 by the chain base method is $I_{12} \times I_{23}$. It will be shown that this is equal to I_{13}.

For simplicity, suppose the G.M. is unweighted. Then, from (6A),

$$\log I_{12} = \frac{1}{n} (\Sigma \log P_2 - \Sigma \log P_1)$$

$$\text{and } \log I_{23} = \frac{1}{n} (\Sigma \log P_3 - \Sigma \log P_2)$$

$$\therefore \log I_{12} + \log I_{23} = \frac{1}{n} (\Sigma \log P_3 - \Sigma \log P_1)$$

$$= \log I_{13}$$

$$\therefore I_{12} \times I_{23} = I_{13} \quad \ldots\ldots\ldots \quad (9)$$

The theorem can easily be generalised to any number of years and to a weighted G.M., provided the weights are kept constant. Normally, however, with the chain base method, the weighting will vary from year to year as the relative importance of the items changes.

The chain base method can also be used when there are marked seasonal variations in production or sales. Each month is " linked up " with the same month of the previous year, and so on to the corresponding month of the base year, each month of which has an index expressed as a percentage of the monthly average for that year. A disadvantage of the chain base method is that if, through any abnormality or defect in the data, the index for any year is unreliable or erratic, the irregularity will be continued in subsequent years, and if there is any bias in the method employed (*e.g.* in the weighting system), the cumulative effect of that bias in a few years' time may become serious. A chain base index is at least as weak as its weakest link.

§ 8.—The Weighting System.

By now the student will realise the importance of the weighting system. If the index covers a large number of items, or a moderately large number of items of almost equal importance, the error involved in using a simple (*i.e.*, unweighted) mean, either A.M. or G.M., will be small, but this rarely happens. Alternatively, if approximate weights can be assigned to the

various items, the result will not be far out, for it is well known that strict accuracy in weighting is not usually necessary. It is important, however, to consider carefully how these weights should be determined, or the results may be peculiar.

For example, it would be absurd to weight an A.M. of price relatives according to the physical weight of the commodities sold. To take an extreme case, if equal weights of gold, silver and copper were sold, no one would think of giving them equal weight in the index. A much better method would be to weight them according to value, *e.g.*, the value of sales in the base year.

§ 9.—The Aggregative Method.

So far two methods of calculation have been considered, with a few variations, (i) the A.M. of price relatives, (ii) the G.M. of price relatives. Other methods involving relatives have been used in the past, *e.g.*, the median of price relatives, but they are not serious rivals of the first two.

There is another method, however, called the Aggregative Method, in which the prices themselves are weighted. An index compiled in this way has the advantages of being readily understood and, very often, easily calculated. It is obtained by taking given quantities of the various commodities and valuing them first at current year's prices, secondly at base year's prices, and dividing the first result by the second.

Let W_1, W_2 .... W_n be the quantities chosen, and as before, let P_1, P_2 .... P_n be the base year's prices

and $P_1', P_2' \ldots P_n'$ the current year's prices. The index is given by

$$I = \frac{W_1 P_1' + W_2 P_2' + \ldots \ldots + W_n P_n'}{W_1 P_1 + W_2 P_2 + \ldots \ldots + W_n P_n}$$

$$= \frac{\Sigma W P'}{\Sigma W P} \quad \ldots \quad \ldots \quad (10)$$

Notice that it is not necessary for the quantities or prices all to be expressed in the same units, but the products W_1P_1, etc., must all be in the same unit before they can be added together and the totals compared.

Quantities commonly selected are those sold in the base year, which may be denoted by $Q_1, Q_2 \ldots \ldots Q_n$. In this case ΣQP is the actual value of the base year's sales, while $\Sigma QP'$ is the amount the same quantities would have realised at current year's prices. Expressing it algebraically,

$$I = \frac{\Sigma Q P'}{\Sigma Q P} \quad \ldots \quad (11)$$

Alternatively, the current year's quantities Q_1', $Q_2' \ldots\ldots Q_n'$ may be taken, giving the formula

$$I = \frac{\Sigma Q' P'}{\Sigma Q' P} \quad \ldots \quad \ldots \ldots (12)$$

By this method the value of the current year's sales is compared with the amount they would have realised at the base year's prices.

If there are many commodities such an index can be compiled by means of a table as shown in the following example, which is purposely simplified, but illustrates the method.

Example 5.

Obtain a price index for 1947 based on 1946 from the following data of sales and prices.

Commodity	Average price 1946	Average price 1947	Sales in 1947	Total Value at 1946 prices	Total Value at 1947 prices
	£/ton	£/ton	Tons	£	£
A	2·43	2·65	3,000	7,290	7,950
B	5·72	5·13	1,200	6,864	6,156
C	10·24	13·09	700	7,168	9,163
D	8·19	8·63	300	2,457	2,589
Total				£23,779	£25,858

$$\text{Price index} = \frac{25{,}858}{23{,}779} \times 100$$
$$= 108{\cdot}7$$

For the purpose of calculating a price index it is not necessary that the values should be expressed in pounds. For instance, if the prices of most commodities are expressed in pence per lb. and the sales in lb. it may be convenient to give the last two columns in thousands of pence.

If there is no reason to prefer "current year weighting" to "base year weighting" or *vice versa*, a combination of the two may be used, the formulae most commonly used being

$$I = \frac{\Sigma\,(Q + Q')\,P'}{\Sigma\,(Q + Q')\,P} \qquad \ldots(13)$$

and

$$I = \sqrt{\frac{(\Sigma\, Q P')\;(\Sigma\, Q' P')}{(\Sigma\, Q P)\;(\Sigma\, Q' P)}} \qquad \ldots(14)$$

Formula (14), which is the G.M. of formulae (11) and (12), is known as Fisher's "Ideal Formula," after its chief exponent, Irving Fisher. Theoretically it is almost perfect, but it requires about twice as much computation as (11) or (12) separately, and there is seldom sufficient difference between the results of these formulae to make the extra labour worth while.

Any of these formulae can be used with either the fixed base or the chain base method.

It may be noted that the aggregative formula is really equivalent to the weighted A.M. of price relatives when the weights are suitably chosen. Formula (11), for instance, can be expressed

$$I = \frac{\Sigma\,(Q\,P)\left(\frac{P'}{P}\right)}{\Sigma\,Q\,P} \qquad (11\text{A})$$

which is the A.M. of the relatives using base year values as weights.

As a general rule, prices should be weighted by quantities and price relatives by values.

§ 10.—Changing the Base Year.

It often happens that when an index has been running for several years it is desired to change the base year in order to make more realistic comparisons. The simplest way is to multiply every value by the same factor, *viz.*, the factor which makes the index of the new base year 100. For example, consider the following index :—

Year	Index
1930	100·0
1931	103·1
1932	105·5
1933	102·6
1934	108·7
1935	119·3
1936	114·8
1937	109·5
1938	125·0
1939	127·4

To change the base year to 1938, each number is multiplied by $\frac{100\cdot0}{125\cdot0}$, *i.e.*, 0·8, and the result, rounding

off each new index to one place of decimals, is as follows :—

Year	Index
1930	80·0
1931	82·5
1932	84·4
1933	82·1
1934	87·0
1935	95·4
1936	91·8
1937	87·6
1938	100·0
1939	102·2

In this case the weighting method used in calculating the index is unchanged, and if the weights represent values or quantities relating to a particular year—the weights base—this will probably differ from the new base year, or, as it is sometimes called, the date base. There is no reason why the weights base and the date base should necessarily be the same, but it usually gives a clearer meaning to the index if they are.

For this reason, and because fixed weights tend to become obsolete, the weighting is often revised at the same time as the base year is changed. Usually, however, only the indices for the period following the new base year are recalculated on the new weighting system, the new indices for previous years being found as described above, by multiplication by a constant factor. It should be noticed that the new indices for recent years will not normally agree exactly with those that would be obtained from the old ones in this way—naturally, or there would be no point in revising the weighting. Thus, in changing the base year to 1938 in the above example, the weighting system would probably be changed from 1938 onwards, giving an index of, perhaps, 101·6 for 1939 instead of 102·2. It must be remembered that the weighting system is

often, to a great extent, arbitrary, and that it is neither necessary nor desirable to apply to back years a weighting system appropriate to a later period.

It is sometimes necessary to link up two or more indices compiled with different base years but overlapping one year or more. The method should now be obvious. Suppose there are two series of index numbers as follows :—

Index A (1930 = 100)		Index B (1938 = 100)	
Year	Index	Year	Index
1930	100 0	1938	100·0
1931	103 1	1939	101·6
1932	105·5	1940	109·7
1933	102·6	1941	115·3
1934	108·7	1942	117 4
1935	119·3	1943	124·8
1936	114·8	1944	123·6
1937	109·5	1945	134·9
1938	125·0	1946	140·5
1939	127·4	1947	147·2

If it is desired to keep 1930 as base year the whole of Index B can be multiplied by 1·25, or if the base year is to be 1938 the whole of Index A can be multiplied by 0·8 as before. Any number of overlapping series of indices can be linked up in the same way.

In a case such as the above, where two different index numbers have been obtained for the same year (in this example 1939) it is quite a common practice to "blend" the two series by taking the A.M. or G.M. of two values for that year. Either the A.M. or the G.M. in this case would give a value of 101·9 (the average of 101·6 and 102·2) for 1939 based on 1938 as 100.

§ 11.—Index Numbers of Quantity.

Indices of quantity, *e.g.*, production, volume of trade, etc., present certain problems of their own,

although the methods of calculation are often similar to those of price indices. For convenience, these problems will be discussed in terms of production.

The first thing is to decide exactly what the index is intended to measure. Obviously not simply the total weight, for it would be absurd to add a ton of limestone to a ton of fertilizer and call the result two tons. The ton of fertilizer is worth far more, and requires more labour to produce, than the ton of limestone. This suggests two possible answers, which will be considered in turn, the first at some length, the second quite briefly.

(i) An index of production may measure the value of output at fixed values per unit of each commodity, *e.g.*, the value of output at base year's prices. This is the sense in which production is measured in the official Interim Index of Industrial Production and the London and Cambridge Economic Service's Index of Production.

Example 6.

In 1935, 300 tons of commodity A and 7,000 gallons of commodity B were produced, worth £5 a ton and 6/- a gallon respectively. In 1946, the output was 200 tons of A and 8,000 gallons of B. Obtain an index of production for 1946, based on 1935.

The total value in 1935 was—

$$£(5 \times 300) + (6 \times 7{,}000)/- \ = \ £3{,}600.$$

The value of the 1946 output, *also at 1935 prices*, was—

$$£(5 \times 200) + (6 \times 8{,}000)/- \ = \ £3{,}400.$$

Hence the required index is—

$$\frac{3{,}400}{3{,}600} \times 100 \ = \ 94{\cdot}4.$$

It is easily seen that this is equivalent to weighting the production relatives by the base year values. In the above example the relatives are $\frac{2}{3}$ and $\frac{8}{7}$ respectively, and the corresponding values in 1935 are £1,500 and £2,100. The index obtained by the weighted A.M. is-

$$\frac{1{,}500 \times \frac{2}{3} + 2{,}100 \times \frac{8}{7}}{1{,}500 + 2{,}100} \times 100$$

$$= \frac{1{,}000 + 2{,}400}{1{,}500 + 2{,}100} \times 100$$

$$= \frac{3{,}400}{3{,}600} \times 100$$

$$= 94{\cdot}4 \text{ as before.}$$

It would, of course, be quite possible to measure the value of production at current prices instead of base year's prices, or to take an average of the two indices so obtained. The analogy with price indices is obvious. **Production figures should be weighted by prices, or production relatives by values.**

An indirect method of obtaining an index of production is to find the price index first and divide it into the total value. Thus if the value of 1946 output, in terms of money, is 120% of that of 1935 output, and the price index for 1946 based on 1935 is 150, the index of production for 1946 based on 1935 is—

$$120 \times \frac{100}{150} = 80.$$

It is often necessary to use this method when the articles in question comprise a large variety of goods of different kinds, grades and sizes, as in the Board of Trade's Import and Export Volume Indices (*see* Chap. X, § 8).

(ii) An index of production may measure the amount of labour effort required to produce the goods. Such an index might be of more interest to the Works Manager than an index based on value. Here it is necessary to estimate, by time study or other means, the average number of man-hours required to produce a given unit of each commodity. Alternatively, the number of man-hours used to produce a unit of product in the base year may be taken. Total output can then be expressed in terms of " standard man-hours " and the index calculated accordingly.

SYNOPSIS TO CHAPTER X.

INDEX NUMBERS IN COMMON USE.

§ 1.—Board of Trade Wholesale Price Index.

2.—"The Statist" Wholesale Price Index Numbers.

3.—"The Economist" Index of Wholesale Prices.

4.—"The Economist" Sensitive Price Index.

5.—Import and Export Price Indices.

6.—Volume and Values of Imports and Exports.

7.—Cost of Living Index Number.

8.—Interim Index of Retail Prices.

9.—Interim Index of Industrial Production.

10.—Index of Industrial Production (London and Cambridge Economic Service).

11.—Rates of Wages Index Number.

12.—Prof. Bowley's Wage-Rate Index Number.

13.—Various Indices of Sales and Stocks.
- (*a*) Index of Retail Sales.
- (*b*) Index of Sales of Independent Retailers.
- (*c*) Trade of Wholesale Textile Houses.
- (*d*) Wholesale Stocks and Sales of Clothing.

CHAPTER X.

INDEX NUMBERS IN COMMON USE.

In this chapter some of the well-known index numbers published by Government Departments and other bodies are described, both on account of their intrinsic importance and to illustrate the methods of compilation. The chapter will be restricted to British index numbers, and even then can only deal with a few of the most important.

It must be remembered that this is being written in 1948, and that there will almost certainly be several additions and amendments to the list by the time it is printed. The student should try to keep this chapter up-to-date himself. He is not expected to learn all the details by heart, but should use it for reference and familiarise himself with the principles and sources of as many index numbers as possible.

§ 1.—Board of Trade Wholesale Price Index.

The present Board of Trade index is based on 1930 and covers 200 price series, some of which relate to commodity-groups with two or more quotations, some to grades of commodities, but most to single commodities. These are derived mainly from weekly or daily quotations published in trade journals, but a few are supplied direct to the Board of Trade by individual firms. The monthly average prices so obtained are compared with those of the corresponding month of the previous year. This is necessary to give a reasonable comparison in the case of commodities such as fruit and vegetables that are not available all the year round.

The geometric mean of the 200 price relatives is used in preference to the arithmetic mean, for the reasons explained in the last chapter. There are eleven main groups of commodities, and the weights assigned to them are roughly proportional to the corresponding values of goods manufactured or produced in the United Kingdom in 1930, according to the Census of Production for that year, together with imports of similar goods passing into consumption without undergoing further manufacturing processes. In assigning these weights, care has been taken to eliminate duplication within groups. The eleven groups and their weights are shown in the table, which is taken from the "Board of Trade Journal" of 8th May, 1948.

Group	No of items	Index No. April 1948	Percentage change compared with Mar 1948	April 1947	Year 1938
I. Cereals	20	175·9	+0 5	+ 2 7	+ 60 1
II. Meat, fish and eggs .	20	138·0	*	+ 9 5	+ 60 6
III. Other food and tobacco . .	28	227·2	+1 6	+19 4	+132·9
Total—Food and tobacco . .	68	182 0	+0·8	+11 4	+ 87·1
IV. Coal	9	293·8	*	+20·3	+138 6
V. Iron and steel . . .	37	234·9	*	+ 7 9	+ 68·9
VI. Non-ferrous metals . .	8	233 6	− 0 1	+ 1·9	+147·5
VII. Cotton	10	321 3	+13 6	+63·1	+284·2
VIII. Wool	11	264·5	− 0·1	+34·7	+160·7
IX. Other textiles	9	165·0	− 0·1	− 0·4	+140·2
X. Chemicals and oils . .	15	189 3	*	+ 9 4	+ 99·9
XI. Miscellaneous	33	267 3	+ 0·3	+36·8	+186·8
Total—Industrial materials and manufactures	132	240·9	+1·0	+20 4	+132·9
Total—All articles	200	219 3	+1·0	+17 2	+116·2
Industrial materials (excluding fuel) :—					
Basic materials . . .	33	296 8	+1·2	+38 1	+219·6
Intermediate products. . .	38	246·9	+0·3	+14·4	+136·3
Manufactured articles	48	213·1	+1·8	+14·3	+ 90·1
Building materials	16	218·7	*	+10·4	+110·1

* No appreciable change.

As the table shows, industrial items other than fuel are also regrouped into " Basic materials," " Intermediate products " and " Manufactured articles," with weights 33, 38, and 48 respectively. There is a further index for " Building materials " compiled from the various quotations for such materials. In this index, however, each commodity is reweighted according to its importance as a building material.

Index numbers for the current and recent months, for all the groups mentioned above, are published once a month in the " Board of Trade Journal," and index numbers for the broader groups are quoted in the " Monthly Digest of Statistics," in which comparison is made with 1938 instead of 1930. Full details of the present index are given in a supplement to the " Board of Trade Journal " for 24th January, 1935.

Previous series were based first on 1913, then on 1924, but the weighting was based on the 1907 Census of Production right up to the introduction of the present series. The number of price-series covered was 150, although the number of quotations was slightly more and gradually increased until the index was replaced by the present one. The grouping was a little different, and the regrouping of industrial items and the index of building materials began with the present series. For further details of the old series, the student is referred to the " Board of Trade Journal" for 20th January, 1921, 24th April, 1930, and 10th December, 1931.

§ 2.—" The Statist " Wholesale Price Index Numbers.

These index numbers are published in great detail each year in the Journal of the Royal Statistical Society. They were begun by Mr. Sauerbeck, and

since 1912 have been compiled by *The Statist.* The base period consists of the eleven years 1867 to 1877, and the index goes back to 1846. There are 45 commodities grouped into six categories as follows :—

Commodity-group.	No. of items
1.—Vegetable food, corn, etc.	8
2.—Animal food	7
3.—Sugar, coffee and tea	4
1-3. Food	19
4.—Minerals	7
5.—Textiles	8
6.—Sundry materials	11
4-6. Materials	26
All groups	45

The method of compilation is very simple. For an annual index, the average annual price is obtained from 52 weekly quotations and expressed as a percentage of the average price for the period 1867-77. The 45 relatives so obtained are shown individually and combined, by means of the simple A.M., to give indices for the various groups. Thus for the second group, animal food, the seven price relatives for 1945 (1867-77 = 100) are 134, 142, 152, 153, 196, 192, 121 : total 1,090. Dividing by 7 gives 156 as the index for the group.

Monthly indices are calculated in the same way, except that the end-of-the-month price is taken instead of the average for the month.

§ 3.—"The Economist" Index of Wholesale Prices.

The Economist Index was started in 1864 and worked back to 1851, with the base period 1845-50. The index was the sum of the price relatives of 22 commodities, the base index being 2,200, not 100 (*see*

Chap. IX, § 1). In 1911 it was revised and based on the period 1901-05, with 44 commodities, but the method of calculation was unchanged (see *The Economist*, 26th August, 1911).

In 1928 the index was completely revised, with base year 1927, the number of commodities being increased to 58, and the final index obtained from the geometric mean of their price relatives (see *The Economist*, 15th December, 1928). The new index was compiled for each month from January, 1924, onwards, and at the time of going to print is still published on the same basis. It now appears fortnightly in *The Economist*, and in the " Records and Statistics Supplement."

Subsidiary index numbers are shown for Cereals and Meat, Other Foods, Textiles, Minerals, and Miscellaneous, as in the table. This is taken from the " Commercial History and Review of 1947, Part I," appearing in the " Records and Statistics Supplement " for 7th February, 1948. The original table also shows each month of 1947, but that part is omitted here for reasons of space. It will be noticed that comparison is still given for 1913 as well as for 1927. In the fortnightly tables, only the figures for the Complete Index are compared with both years.

"THE ECONOMIST" INDEX OF WHOLESALE PRICES

Commodity	Unit of Measurement	Average					
		1929	1939	1944	1945	1946	1947
Cereals and Meat	1913=100	135·6	103·1	151·1	151·4	153·0	153·1
Other Food ..	,,	136·8	104·5	171·0	171·0	173·3	218·1
Textiles ..	,,	141·9	90·5	155·1	161·2	175·2	216 5
Minerals ..	,,	116·2	116·3	172·7	183·9	221·6	272·6
Miscellaneous ..	,,	112·7	99·8	160·4	160·3	163·3	184·3
Complete Index	,,	**127·2**	**102·1**	**161·5**	**164·2**	**174·2**	**201·5**
Cereals and Meat	1927=100	96·9	73·8	108·1	108·2	109·3	109·4
Other Food ..	,,	85·8	65·6	107 3	107·3	108·7	136·9
Textiles ..	,,	92·2	58·8	100·9	104 7	113·8	140 6
Minerals ..	,,	95·6	95·6	141·9	151·3	182·3	224·2
Miscellaneous ..	,,	90·5	80·8	128·9	128·8	131·1	148·0
Complete Index	,,	**92·4**	**74·2**	**117·1**	**119·3**	**126·8**	**146·4**

§ 4.—" The Economist " Sensitive Price Index.

The " Sensitive Price Index " was introduced in October, 1940, to register changes from day to day in the state of markets for primary products (see *The Economist,* 26th October, 1940). The base year is 1935, and the index is calculated for the end of each month, beginning at January, 1913, and for every day since April, 1938.

The index covers only ten commodities, (*a*) five subject to discontinuous production, *viz.*, wheat, maize, sugar, cocoa and cotton, and (*b*) five subject to continuous production, *viz.*, rubber and the non-ferrous metals copper, lead, tin and zinc. The index is the simple geometric mean of the ten price relatives. Subsidiary indices are also compiled separately for groups (*a*) and (*b*), which are described in the tables as Crops and Raw Materials respectively.

The index is published each week in the " Records and Statistics Supplement " to *The Economist,* from which the following table is taken :—

" THE ECONOMIST " SENSITIVE PRICE INDEX
(1935 = 100)

	May 10, 1947	May 4, 1948	May 5, 1948	May 6, 1948	May 7, 1948	May 10, 1948	May 11, 1948
Crops	190·7	226·5	226·5	226·5	226·5	226·5	226·5
Raw materials	360·5	369·2	369·6	370·1	370·1	369·6	370·3
Total Index ..	262·1	289·2	289·4	289·6	289·6	289·4	289·6

As *The Economist* said in October, 1940, " the value of the new price indices has been temporarily (sic !) impaired by the imposition of price control."

§ 5.—Import and Export Price Indices.

Since the Second World War, there has been much discussion about "terms of trade," *i.e.*, the rate at which, broadly speaking, exports are exchanged for imports. It is often said that the terms of trade are moving against the United Kingdom, meaning that prices of imports are rising more rapidly than prices of exports, and hence a given "unit" of exports obtains an ever-diminishing quantity of imports in exchange. The Board of Trade's *Indices of Import and Export Prices* are intended to show the change in prices of imports and exports in the current period as compared with 1938. They appear monthly in the "Board of Trade Journal" and, in a different form, in the "Monthly Digest of Statistics."

The method of construction is described in the "Board of Trade Journal" for 13th April, 1946, and is somewhat complicated—in the writer's opinion unnecessarily so. For convenience the following account will refer to imports. At any particular time, the import price index for each month of the previous year and the current year up-to-date is calculated on an aggregative formula with fixed weights, representing the estimated imports for the current year. Thus 13 to 24 months, according to the time of the year, are shown on a comparable basis in the "Monthly Digest of Statistics," as in the table overleaf.

For each month from January, 1946, to July, 1947, the estimated imports for 1947 have been valued at (i) the average import prices for the month and (ii) the average import prices of 1938, and the first result divided by the second.

		Imports (1938=100)			
		Total	Food, drink and tobacco	Raw materials and articles mainly unmanufactured	Articles wholly or mainly manufactured
1939	December	130	123	137	142
1940	December	154	146	163	160
1941	December	164	152	180	173
1942	December	179	167	200	188
1943	December	188	172	217	192
1944	December	195	181	221	190
1945	December	195	195	214	172
1946	January	198	199	216	176
	February	199	200	216	177
	March	199	200	216	177
	April	200	200	217	181
	May	201	201	219	180
	June	203	202	221	182
	July	205	203	222	188
	August	206	204	223	189
	September	208	205	226	195
	October	215	215	228	198
	November	217	217	231	202
	December	219	216	235	206
1947	January	223	222	237	208
	February	226	225	242	210
	March	229	226	246	216
	April	234	229	254	223
	May	241	232	268	232
	June	245	236	271	237
	July	252	241	282	244

For the previous years, however, the weights used are the approximate imports *during those years*, and the index shown for, say, December, 1943, is the geometric mean of two index numbers for that month, one calculated on 1943 weighting and the other on 1944

weighting. Consequently, the figures for December of each year from 1939 to 1945, though reflecting the general trend of import prices, are not strictly comparable with the succeeding monthly figures. Incidentally, it may be noticed that the indices for December, 1938, would not necessarily be 100 as their omission might suggest.

In 1948, the months shown would be January, 1947, to the current month of 1948, and each month of 1947 would be recalculated with 1948 weighting. Thus every month of each year is eventually worked out on two weighting systems, that for the current year and that for the following year.

The headings used in the indices have been carefully selected as representative of the whole and not subject to erratic fluctuations in average value on account of changes in size, grade, etc., *i.e.*, the average value must behave like a true price. Thus, lathes have been excluded because their make-up varies greatly from month to month and an average value would be meaningless. The numbers of headings in each main group, and the percentage of total declared value covered by the indices in 1946 are as follows :—

Index	No. of Headings	Proportion of total declared value covered by Index
Imports :		%
Food, drink and tobacco	48	72
Raw materials ..	45	73
Manufactured articles	33	60
Total	126	70
Exports :		
Food, drink and tobacco	8	39
Raw materials ..	8	35
Manufactured articles—		
metal goods ..	41	53
textiles	32	47
other	32	44
Total	121	48

§ 6.—Volume and Value of Imports and Exports.

Indices of volume of imports and exports are published in the "Board of Trade Journal" and, in less detail, in the "Monthly Digest of Statistics." In the former they are accompanied by *Indices of Average Values*, which must not be confused with the Import and Export Price Indices. There are separate indices for imports, retained imports, exports, and re-exports. The accompanying table for imports, from the "Board of Trade Journal" of 8th May, 1948, is typical.

Period	Trade as declared	Trade on basis of 1938 average values	Index Numbers	
			Average values	Volume
	£ million	£ million	(1938=100)	
		TOTAL IMPORTS		
1938	919·5	919·5	100	100
1946 . . .	1301 0	611·9	211	67
1947 1st Quarter .	364·0	152·9	238	67
2nd Quarter .	455·5	173·9	262	76
3rd Quarter	517·4	194·2	266	84
4th Quarter	450·9	170 7	264	74
1948 1st Quarter . .	487·5	178·5	273	78

Here the imports for 1946 and each subsequent quarter have been revalued at the average values of imports in 1938, as shown in Accounts of Trade and Navigation of the United Kingdom, including estimates where necessary. Thus at 1938 average values, the imports for the 1st quarter of 1947 would have been £152·9m. and imports for the average quarter of 1938 were £229·9m. ($\frac{1}{4}$ of £919·5m.). The import volume index for the period is therefore

$$\frac{152{\cdot}9}{229{\cdot}9} \times 100 \quad = 67 \text{ to the nearest integer.}$$

The actual value of imports for the quarter was £364·0m., so the index of average values is

$$\frac{364{\cdot}0}{152{\cdot}9} \times 100 \quad = 238 \text{ to the nearest integer.}$$

The first calculation shows that the physical volume of goods imported in the first quarter of 1947 was about 67% of that for the average quarter of 1938. The second shows that the goods actually imported in the same quarter cost 238% of what they would have cost in 1938. Since the pattern of trade, particularly imports, varies from quarter to quarter, it would be quite possible for the Index of Average Values to fall even if prices rose. The Import and Export Price Indices, therefore, are a better indication of movements in the terms of trade from month to month or from quarter to quarter, although over a longer period it is necessary to use the Indices of Average Values.

The Indices of Volume and Average Values have been published for many years and on various base years. In 1937, the base year was changed from 1930 to 1935 (*see* "Board of Trade Journal," 4th March, 1937). During the 1939-45 War publication was suspended, and after the War the indices were based on 1938 (*see* "Board of Trade Journal," 4th May, 1946). It was stated at the time that "five years is about the normal maximum length of time that a series of index numbers of this kind should be allowed to run without a change of base year, and a change will be made to a post-war base year as soon as circumstances permit." At the time of writing, however, the base year is still 1938.

§ 7.—Cost of Living Index Number.

This Index Number was discontinued in June, 1947, but is still of considerable interest and historic importance. It was started during the First World War in order to measure the cost of living of the average working-class family, taking July, 1914 as base, and was published each month in the "Ministry of Labour

Gazette." It did not attempt or claim to represent the cost of living of the middle or upper classes, which might be quite different ; nor did it take into account any changes in the standard of living or the distribution of expenditure since its inception. Its purpose was simply to measure the cost of maintaining unchanged the standard of living prevailing among working-class households in July, 1914.

The items included in the index were grouped into five broad groups, *viz.*, (1) Food, (2) Rent, (3) Clothing, (4) Fuel and Light, and (5) Other Items. Current retail prices, rents, etc., were obtained each month from representative retailers, house agents, local authorities, coal merchants, etc., and compared with the corresponding prices of July, 1914. The price relatives were then combined, using weighted arithmetic means, into index numbers for each group and for all groups together. The weights were derived from the results of a survey taken in 1904 (!) by the Board of Trade, in which the budgets of less than 2,000 working-class families were collected and analysed. For instance, the average family income was 36s. 10d., and of that 22s. 6d., or about 60% was spent on food, consequently the weight attached to food is 60 out of a total of 100.

The weights of the five main groups are as follows :

Group.	Weight.*
Food	60
Rent	16
Clothing	12
Fuel and Light	8
Other Items	4
Total	100

*The Ministry of Labour divides all these percentages by 8, making a total of 12½, but there seems to be no reason to introduce fractions.

It was thought that there had been little change in the standard of living between 1904 and 1914, so these

weights are roughly proportional to the estimated expenditure in 1914. A very clear and detailed account of the index is given in "The Cost of Living Index Number, Method of Compilation," a white paper issued by H.M. Stationery Office, price 2d.

§ 8.—Interim Index of Retail Prices.

Long before the 1939-45 War it was realised that the official Cost of Living Index Number was quite out-of-date, and in 1937-38 a new enquiry was made, in which family budgets were collected from over 10,000 working-class households all over the country. Unfortunately the results were not available until war had broken out, and the old Index Number continued in use, as already stated, till the 17th June, 1947, the last date for which it was published.

It was replaced by an "Interim Index of Retail Prices," which at the time of writing is being published each month in the "Ministry of Labour Gazette" and the "Monthly Digest of Statistics," and is intended to continue until a new survey can be carried out and a more permanent index worked out. The Interim Index shows the changes since the 17th June, 1947, of the retail prices of the things in which the working-class households of 1937-38 were interested. It does not show the rise in the cost of living since 1937-38. Indeed, it makes no claim to measure that elusive concept, the cost of living—it is simply an index of retail prices, although as a cost of living index it would be a great improvement on the old one.

The fact that it is based on the 1937-38 pattern of expenditure means that many articles are included in the new index that did not appear in the old, such as fresh fruit and vegetables, jam and marmalade, cake and biscuits, pork, lamb, alcoholic drinks, furniture, electricity, wireless sets, prams, slippers, books, car

licences and petrol, and entertainments. A detailed comparison of the items covered in the two index numbers will be found in a white paper entitled "Interim Index of Retail Prices—A Short Explanatory Note," price 2d. A much more detailed account of the new index is given in "Supplement No. 2, January, 1948," to the "Industrial Relations Handbook, 1944," price 6d. Both booklets are published by H.M. Stationery Office.

As might be expected, the weighting system of the new index is very different from that of the old. The weights of the main groups are as follows :—

	Group.	Weight.
i.	Food	348
ii.	Rent and Rates	88
iii.	Clothing	97
iv.	Fuel and Light	65
v.	Household Durable Goods	71
vi.	Miscellaneous Goods	35
vii.	Services	79
viii.	Drink and Tobacco	217
	Total	1,000

As in the old Cost of Living Index Number, the index for "All Items" and the index numbers for each main group are weighted arithmetic means of the price relatives of individual items.

§ 9.—Interim Index of Industrial Production.

This index is prepared by the C.S.O. (Central Statistical Office) in collaboration with the Board of Trade and other Government Departments, and published each month in the "Monthly Digest of Statistics" and the "Board of Trade Journal." The original accounts of it appeared in the "Monthly Digest of Statistics" for February, 1948, and the "Board of Trade Journal" of 13th March, 1948. It is to be superseded by a more permanent index when the results of the 1948 Census of Production are available.

The base year is 1946, so that the level of production each month, beginning with January, 1946, is expressed as a percentage of the average monthly production in 1946. Adjustments are made for variations in the number of days (excluding Sundays) in the month, but not for holidays or seasonal factors of any kind. The index covers mining, manufacturing, building and public utilities, but excludes agriculture, transport and distribution, commerce and services.

Index numbers for groups of industries, and for industry as a whole, have been obtained as weighted arithmetic averages of relatives for about 400 items. At the time of writing, the main industrial groups and the weights allocated to them are as shown in the table, industries being classified according to the Standard Industrial Classification.

Standard Industrial Classification Order Number	Industrial Group	Weight
ii.	Mining and quarrying	78
iii.	Treatment of non-metalliferous mining products other than coal—	
	China and earthenware	4
	Glass	6
	Bricks, cement and other non-metalliferous mining manufactures	21
iv.	Chemicals and allied trades	65
v.	Metal manufactures—ferrous	38
	non-ferrous	18
vi.	Engineering, shipbuilding and electrical goods	186
vii.	Vehicles	95
viii.	Metal goods not elsewhere specified	36
ix.	Precision instruments, jewellery, etc	8
x.	Textiles	55
xi.	Leather, leather goods and fur	6
xii.	Clothing	38
xiii.	Food, drink and tobacco—food	60
	drink and tobacco	67
xiv.	Manufactures of wood and cork	25
xv.	Paper and printing	39
xvi	Other manufacturing industries	19
xvii.	Building and contracting	92
xviii.	Gas, electricity and water	64
	All industries	1.000

These weights are roughly proportional to the estimated net output of each industry in 1946. In the absence of a Census of Production for that year, these estimates have been based on (i) the net output in 1935 as shown by the 1935 Census, (ii) estimates of the net output of small firms not covered by the Census, and (iii) estimated changes in the total wage-bill of each industry between 1935 and 1946.

A brief reference should be made here to the Board of Trade's original Index of Production, started in 1928 with 1924 as base year, and published only once a quarter. When the results of the 1930 Census of Production became available the base year was changed to 1930 and slight modifications made, such as the addition of a new group, Building Materials and Building. In this form it continued till August, 1939. The methods of construction were similar to those employed in the Interim Index. Full details are given in the "Board of Trade Journal," 26th July, 1928, and 28th March, 1935.

§ 10.—Index of Industrial Production (London and Cambridge Economic Service).

The chief difference between this index of production and the official Interim Index is that the latter includes finished munitions and repair work, whereas the L. & C. Index excludes them. Otherwise the two indices are remarkably similar in general principles and methods of construction. They are both based on 1946 and compiled by weighted arithmetic means of relatives, with weights proportional to estimated net output in 1946. The grouping of industries, however, is different. The following table shows the groups and their weighting in the L. & C. Index as given in the Bulletin for February, 1948:—

Industrial Group	Weight
Textiles	76
Clothing and Leather	51
Metal Production	61
Shipbuilding and Repairing—A (see below)	23
B (see below)	23
Motors, Cycles and Aircraft	31
Industrial Machinery and Equipment	124
Other Metal-Using	116
Food, Drink and Tobacco	120
Chemical and Allied Trades	66
Building, Building Materials & Furniture—A	101
B	111
Fuel and Power	143
Paper and Printing	50
Sundry Trades	38
Total—A	1,000
B	1,010

In two groups, Shipbuilding and Repairing, and Building, Building Materials and Furniture, alternative indices are given. The "A" series is based on the quantity of goods delivered, while the "B" series takes into account changes in work-in-progress. Thus, the "A" index for building would reflect the number of houses completed in the period, whereas the "B" index would measure the amount of work done even if no houses were actually completed in the period. For most industries lack of information necessitates indices based on deliveries rather than work done.

Four different types of "indicators" have been used. In about two-thirds of the industries the indicator used is the physical quantity of goods produced, *e.g.*, tons, barrels, etc. In many cases, however, this method is not suitable owing to absence of data or differences in grade or other forms of heterogeneity, and other methods used are (i) the value of goods produced, corrected for price changes (*see* Chap. IX, § 11), (ii) the quantity of the principal raw materials consumed, and (iii) numbers employed.

A detailed account of the index is given in the Bulletin of the London and Cambridge Economic Service for February, 1948. In the Bulletin for May, 1948, an attempt is made to compare production in 1935 with that of 1946 and later years, but lack of adequate data for comparing years so far apart makes it extremely difficult to obtain more than approximate results.

§ 11.—Rates of Wages Index Number.

For many years the Ministry of Labour has compiled an index of wage rates, derived from the standard rates of wages fixed by collective agreements or statutory orders. The present index is published monthly in the " Ministry of Labour Gazette " and the " Monthly Digest of Statistics." It includes manual workers and shop assistants, but excludes clerical, technical and administrative workers, and covers Agriculture, Mining, Manufacturing Industries, Building, Civil Engineering, Transport, Public Utility Services, and Distributive, Catering, etc., Trades. It is based on the 30th June, 1947.

Where practicable, account is taken separately of rates for skilled and unskilled workers, men and women, boys and girls, and movements in piece-rates as well as in time-rates. Wage-relatives are calculated for men, women, juveniles and all workers combined, in each industry, and weighted according to the appropriate wage-bills for 1946. In other words, the index for the current month for men, say, is obtained by comparing with the actual wage-bill for men in 1946, the wage-bill as it would be for the same distribution of men at present rates of wages.

The following table is taken from the " Ministry of Labour Gazette " for May, 1948 :—

Date (end of month)	Men	Women	Juveniles	All Workers
1947				
June	100	100	100	100
July	100	100	100	100
August	101	101	102	101
September	101	101	102	101
October	101	103	103	102
November	103	103	105	103
December	103	103	106	103
1948				
January	104	104	106	104
February	104	105	106	104
March	104	106	107	105
April ..	105	107	107	105

An account of the present index appears in the "Ministry of Labour Gazette" for February, 1948. The previous index was based on the beginning of September, 1939, and the one before that on the average for 1924, the principles and methods of compilation being very similar to those of the present index.

The Index of Wage Rates must not be confused with an index of earnings, which would take into account movements between industries, variations in the proportions of skilled and unskilled workers, the amount of overtime worked or bonuses paid, the numbers paid by piece-rates, and possibly such items as holidays with pay, pension schemes, etc. Although the Ministry of Labour does not publish an index of earnings as such, it does show every half-year the percentage increases in earnings in different industries, most of these factors being taken into account.

§ 12.—Prof. Bowley's Wage-Rate Index Number.

This index of weekly wage-rates is compiled monthly by Prof. A. L. Bowley and issued quarterly in the Bulletin of the London and Cambridge Economic Service. It first appeared in January, 1923, based on 1914, and consisted of a simple arithmetic mean of relatives for 11 groups of occupations. In 1929 it was

revised from 1924 onwards, based on December, 1924, and extended to include 20 occupation-groups (*see* table below). These groups were allotted weights based on the weekly wage-bills in 1924, with certain adjustments to give the correct proportions to men and women, and to the three groups, Coal, Agriculture and Other Industries.

The 20 occupation-groups and the weights given to them are shown below. Nos. 1-12 are based on time-rates for a full normal working week, No. 14 equally on time and piece-rates, Nos. 13 and 15 on piece-rates modified in recent years by other factors, and Nos. 16-20 on women's minimum time-rates. Of the last group, No. 16 is based on the weekly rate and the others on hourly rates multiplied by the number of hours in the normal working week, reduced since the war.

Occupation-group	Weight
1.—Compositors	3
2.—Bricklayers	11
3.—Bricklayers' labourers	4
4.—Engineer fitters	12
5.—Engineer labourers	7
6.—Dock labourers	3
7.—Railwaymen	11
8.—Tram drivers and conductors	3
9.—Lorry drivers	3
10.—Local Authority labourers	4
11.—Agricultural labourers	4
12.—Shipbuilders	4
13.—Cotton operatives	10
14.—Wool operatives	5
15.—Coal-miners	10
16.—Women employed in boot-making	1
17.—Women employed in tailoring	2
18.—Women employed in shirtmaking	1
19.—Women employed in confectionery	1
20.—Women employed in tobacco	1
Total	100

Further details will be found in Special Memorandum No. 28 of the London and Cambridge Economic

Service, entitled " A New Index Number of Wages," by A. L. Bowley, January, 1929, or in the Bulletin for January, 1944.

§ 13.—Various Indices of Sales and Stocks.

The " Board of Trade Journal " publishes every month several indices of retail sales, wholesale textile sales and stocks, etc., which, for reasons of space, cannot be described here in great detail. Unlike most official statistics, they are all compiled from information supplied by voluntary contributors.

(a) Index of Retail Sales.

This index was compiled for many years by the Bank of England in collaboration with various trade associations. It covers the trade of a sample of department stores, multiple shops, retail co-operative societies, and a few independent retailers. Since 1947, the Board of Trade has taken over this work from the Bank of England. Originally, comparisons were made of average daily sales value, but the figures are now obtained as far as possible for 4 or 5-weekly periods and the index compiled on a weekly basis. Sales are valued at retail selling value and include purchase tax. Separate indices are given for 7 districts and various categories of goods. The " Monthly Digest of Statistics " gives indices for three broad groups, Food and Perishables, Apparel, and Household Goods. The present base year is 1942.

Further details will be found in the " Board of Trade Journal " for 15th March, 1934, 9th February, 1939, 1st April, 1944, and 19th April, 1947.

(b) Index of Sales of Independent Retailers.

This is a comparatively new index of sales reported direct to the Board of Trade by a selected sample of independent retailers, originally confined to furniture,

chemists' goods and mens' wear, but since extended to various other non-food trades. Contributors supply figures for periods of 4 or 5 weeks, and each month is linked up to the corresponding month of the previous year to give an index of weekly sales, the present base year being 1947. Full details are given in the "Board of Trade Journal" for 19th October, 1946, 24th May, 1947, and 8th November, 1947.

(c) Trade of Wholesale Textile Houses.

Index numbers under this heading are compiled by the Wholesale Textile Association, in collaboration with the Bank of England, from figures supplied by members of the Association. Separate indices are calculated for Women's and Children's Wear, Men's and Boys' Wear, and Dress Materials. Sales are valued at wholesale selling value, excluding purchase tax, while stocks are valued at cost. They are based respectively on the monthly average sales and stocks for the base year, which at the time of writing is 1942.

For further particulars see the "Board of Trade Journal" for 25th May, 1939 and 2nd September, 1944.

(d) Wholesale Stocks and Sales of Clothing.

The indices under this heading differ from those described in section (*c*) chiefly in representing changes in the number of garments, not in their value. They are given in great detail, both stocks and sales being based on average monthly sales in 1947. Combined indices of purchases, sales and stocks, measured in the coupon value of garments and weighted according to the coupon value of sales in 1947, are given for Hosiery, Women's, Girls' and Infants' Made-up Garments, Men's and Boys' Made-up Garments, and All Garments.

Accounts of these index numbers appeared in the "Board of Trade Journal" for 5th April, 1947, and 22nd May, 1948.

SYNOPSIS TO CHAPTER XI.

CORRELATION.

§ 1 —Definition of Correlation.

2.—Correlation of two Factors.

3.—Co-efficient of Correlation.

(a) For Long Term Fluctuations.

(b) For Short Term Fluctuations.

(c) The Probability of Error.

4 —Ratio of Variation

CHAPTER XI.

CORRELATION.

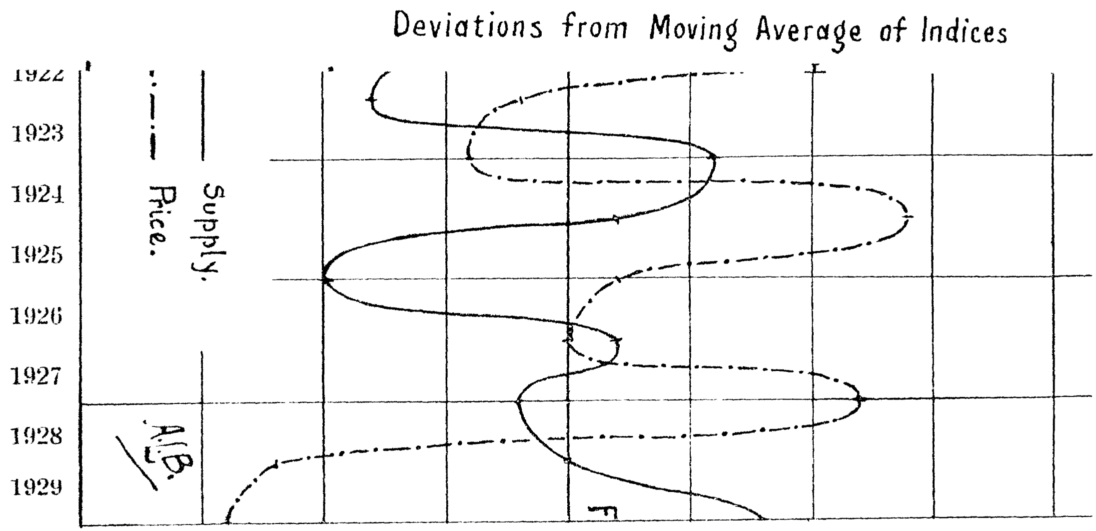
So far we have studied the methods by which comparison between groups, classes or series of data can be carried out, but it must be recognised that we must also study any relationship which exists between various factors. It has been stated previously that *all other things being equal profits should increase with the amount of capital invested in a business.* In such a case there is obviously direct connection. Similarly there must be some connection between the Turnover of a business and the Profits ; and the Output of a factory should be closely related to the number of Man-hours worked. The economic law of supply and demand indicates that if no other disturbing factors are present the price of a commodity declines as the supply increases, and advances when the supply diminishes.

§ 1.—Definition of Correlation.

Whenever some definite connection exists between two or more groups, classes or series of data, there is said to be correlation. It does not matter whether the data in one section changes in the same or in the reverse direction to that in the other, so long as a movement in sympathy is apparent. In commercial work it is very important that we should know that correlation exists, and to what extent it is present.

§ 2.—Correlation of Two Factors.

If we take two such factors as the supply and price of a commodity (or commodities) and desire to ascertain



ıd Price

facing p. 288

the effect which the supply has upon price, we must consider the data from two aspects. We must, firstly, eliminate the short-term fluctuations and consider the trend of both price and supply; secondly, we must eliminate the long-term fluctuations in order to consider the effect that abnormal conditions in one factor have upon the other at the time the condition rules.

We have already seen that we can eliminate the short-term fluctuations of any single series of data by plotting the moving average line, and can then measure the effect of abnormal conditions by comparing the curve of actual figures with the curve of the moving average.

If reference be made to Diagram No. 8, the fluctuations from the moving average line or trend can be very clearly seen, and the difference can easily be measured. As we are already aware, these short-term fluctuations have no marked effect upon the trend, as may be seen from the regularity of the upward movement in the moving average line in the diagram mentioned, though the actual figures were subject to wide variations.

When comparison between two sets of data is necessary, the data to be compared should be placed upon the same chart, but as different units and types are being used, great care has to be exercised in plotting the data if the true comparison is to be found. This difficulty can be eliminated by using Index Numbers for each of the factors, as they will then be reduced to similar terms and can be plotted on the same scale. Columns 1 and 4 of Table Y (on next page) show the Index Numbers of the Supply and Price respectively of any commodity. The figures are entirely hypothetical, and have been compiled for the purpose of

illustration only. Columns 2 and 5 show the three-yearly average of the Indices so gathered. The figures in these four columns have been plotted in Diagram No. 23, where the curve A shows the actual

TABLE Y.

Year	Supply.			Price.		
	1	2	3	4	5	6
	Index Numbers	Moving Average of Indices.	Deviation from Moving Average	Index Numbers.	Moving Average of Indices.	Deviation from Moving Average.
1900	106	—	—	86	—	—
1901	104	—	—	90	—	—
1902	100	103	— 3	94	90	+ 4
1903	102	102	0	89	91	— 2
1904	101	101	0	93	92	+ 1
1905	100	101	— 1	97	93	+ 4
1906	99	100	— 1	98	96	+ 2
1907	104	101	+ 3	91	95	— 4
1908	103	102	+ 1	90	93	— 3
1909	96	101	— 5	102	94	+ 8
1910	101	100	+ 1	93	95	— 2
1911	100	99	+ 1	93	96	— 3
1912	93	98	— 5	102	96	+ 8
1913	95	96	— 1	99	98	+ 1
1914	103	97	+ 6	99	100	— 1
1915	90	96	— 6	105	101	+ 4
1916	98	97	+ 1	100	101	— 1
1917	106	98	+ 8	95	100	— 5
1918	93	97	— 4	108	101	+ 7
1919	89	96	— 7	103	102	+ 1
1920	103	95	+ 8	98	103	— 5
1921	93	95	— 2	111	104	+ 7
1922	92	96	— 4	103	104	— 1
1923	97	94	+ 3	104	106	— 2
1924	96	95	+ 1	114	107	+ 7
1925	89	94	— 5	110	109	+ 1
1926	94	93	+ 1	112	112	0
1927	90	91	— 1	120	114	+ 6
1928	92	92	0	107	113	— 6
1929	97	93	+ 4	103	110	— 7

Indices of the Supply of the Commodity, and curve *a* the moving average of the indices or the trend of the supply. The other curves marked B and *b* show the Indices and the averages respectively of the price.

The curves *a* and *b* can be compared with a view to discovering if there is any relation existing between the two over the period under review. If the curves move together in the same or opposite direction, then some degree of relationship apparently exists in the trends. It will be seen in this case that even after eliminating the short-term fluctuations the line showing the trend of supply moves in the opposite direction to the curve for the moving averages of the indices of price. There are minor differences, as is only to be expected, but the general tendency of the curves is in opposite directions, thus indicating that price recedes where supply advances, and *vice versa.* If we now eliminate the long-term fluctuations and compare only the short-term variations, we may obtain other information of great value. In columns 3 and 6 are tabulated the deviations which the index for the year shows from the moving average or trend, and these deviations have been plotted in Diagram No. 24. The zero line (marked EF) may be regarded as the straightened moving average lines in the previous diagram, and the curves show how the indices for supply and price deviate from this line. It will at once be seen that in nearly all cases when supply is low, prices are high, and *vice versa,* but in 1922 and 1925 the deviations appear to be in the same direction. The deviations vary considerably in magnitude, *e.g.*, in 1909 and again in 1912 an equal fall in the supply was accompanied by rises in price which were identical in each case, but in 1915, though the supply was even shorter, the price did not rise so sharply as in the other two years. In several years the highest rise in one exactly corresponded in time with the lowest of the others (*c.f.* 1907, 1909, 1912, 1915, 1929), but in a few cases there was a "lag" in one of the factors. In 1918, for example, the price deviation was high, and

then fell away in the following year; but though supply fell in 1918, it was lower still in 1919. Again, in 1921 the price deviation was high and that for supply low, but in 1922 both price and supply receded. These circumstances indicate that other factors, such as speculation, were influencing the price, but it is clear that the supply is the main factor regulating price. Knowing this is so in any particular case, the astute business man in possession of information as to the likelihood of supplies being short would anticipate the rise in price by buying early. This is where valuable information is given by commercial newspapers in supplying estimates of crops, output, etc.

§ 3.—Co-Efficient of Correlation.

(a) For Long Term Fluctuations.

While the method outlined above is satisfactory for most purposes, particularly for commercial data and with a view of ascertaining if correlation is present in the factors under review, yet it affords no means of measuring the degree of correlation which exists in the data. It is necessary therefore that a co-efficient of correlation be prepared in order to render comparison possible. To do this we class the two factors we are investigating, the one as the **Subject** and the other as the **Relative.** The "Subject" is the factor we desire to use as the standard, while the "Relative" is the variable we are desirous of comparing with the subject, *e.g.*, in the foregoing example the Supply of the commodity is the Subject, and the price may be regarded as the Relative.

When it is desired to measure the correlation present in long-term fluctuations the most satisfactory co-efficient is that calculated by Karl Pearson for measuring biological correlation. It (r) is obtained as follows: *The deviation of each "subject" item from*

the Arithmetic Average of all the Subject items (x) is multiplied by the deviation of each "Relative" item from the Arithmetic Average of all the Relative items (y), and the results so obtained are summated and divided by the number of items under review (n), multiplied by the Standard Deviation of the Subject items (σ_1) and by the Standard Deviation of the Relative items (σ_2) i.e.,

$$\therefore r = \frac{\Sigma (x\,y)}{n\ \sigma_1\ \sigma_2}$$

Example :

SUBJECT.			RELATIVE.			
Capital in Hundreds of £'s.	Deviation from Average.	Square of Deviation.	Profit in Hundreds of £'s	Deviation from Average	Square of Deviation.	(x) × (y)
(a).	(x)	(x²).	(d)	(y)	(y²)	(g)
10	− 45	2025	2	− 11	121	+ 495
20	− 35	1225	4	− 9	81	+ 315
30	− 25	625	8	− 5	25	+ 125
40	− 15	225	5	− 8	64	+ 120
50	− 5	25	10	− 3	9	+ 15
60	+ 5	25	15	+ 2	4	+ 10
70	+15	225	14	+ 1	1	+ 15
80	+25	625	20	+ 7	49	+ 175
90	+35	1225	22	+ 9	81	+ 315
100	+45	2025	30	+17	289	+ 765
Aver. 55		Σx^2 8250	Aver. 13		Σy^2 724	$\Sigma(xy)$+2350

Then the Standard Deviation of the Subject items σ_1 =

$$\sqrt{\frac{8250}{10}} = 28{\cdot}72$$

and the Standard of Deviation the Relative items σ_2 =

$$\sqrt{\frac{724}{10}} = 8{\cdot}507$$

therefore the Co-efficient of correlation =

$$\frac{2350}{10 \times 28{\cdot}72 \times 8{\cdot}507} = \frac{2350}{2443} = 0{\cdot}9618$$

It must be borne in mind that in calculating both x and y due attention must be given to the signs preceding the deviations so that $+x$ and $-y$ will give $-xy$ and the sum is the algebraic sum.

When there is perfect correlation between the factors under consideration it is represented by the symbol 1. If the correlation is direct, *i.e.*, the deviations of both factors are in the same direction as is the case in the example just given, the positive sign (+) is prefixed to the symbol; while if the movement in the Relative is in the opposite direction to that in the Subject the negative sign (–) is used as a prefix. The symbol O indicates that no correlation exists. In the above example therefore the correlation would be expressed as + ·9618.

Note.—When the deviations in the Subject and Relative are in the same direction the deviations are said to be "concurrent."

(b) For Short Term Fluctuations.

Occasionally it is found necessary to study the correlationship existing in short term fluctuations, and when this is required an adaptation of the above co-efficient of correlation is used. Instead of using the deviations of the items of the Relative and Subject from the Arithmetic Average of the items in each case we calculate the deviations from the Trend. In such a case the Moving Average of the Index Number of the two factors is calculated, and the deviations of such figures from the average of the Indices is the measure of deviation which is used for obtaining the Standard Deviation in each of the cases. With this difference the formula and the method of calculation are exactly the same.

(c) The Probability of Error.

In considering any question of correlation we must recognise the fact that if only a few representative items are used in calculating the co-efficient of correlation chance may result in an apparent connection

being present, whereas if a larger number of examples were taken it would be proved to be non-existent. In order to guard against false assumptions being made it is usual to calculate the probable error in the co-efficient, and use it in the same way and for the same reason as the probable error which is present in an estimate. The formula for calculating the probable error in the co-efficient of correlation is as follows :—*Subtract the square of the co-efficient of correlation from unity, multiply by* 0·6745, *and divide the result by the square root of the number of items used in obtaining the co-efficient.* In the example given below, therefore, the probable error is :—

$$\frac{\cdot 6745\ (1 - \cdot 9618^2)}{\sqrt{10}} = \frac{\cdot 6745\ (1 - \cdot 9251)}{3\cdot 162}$$

$$= \frac{\cdot 6745 \times \cdot 0749}{3\cdot 162} = \frac{\cdot 05053}{3\cdot 162} = \cdot 01598.$$

The co-efficient of correlation would therefore be written 0·9618 ± 0·01598.

If the co-efficient is less than the probable error obtained in this manner it is apparent that there can be no correlation between the factors, but when the co-efficient is larger than the probable error, correlation can be assumed. It is not advisable, however, to rely upon its presence unless the co-efficient is several times larger than the probable error. Obviously therefore the larger the number of times the co-efficient exceeds the probable error, the greater the degree of correlation.

§ 4.—Ratio of Variation.

As will be seen from diagram No. 24 the movements of the variables while generally corresponding as to

time do not necessarily agree in the degree to which they move. In the year 1912, for example, the price movement was nearly twice as great as that of supply. It becomes imperative, therefore, to know what is the average ratio existing between the proportional or percentage deviation of the two curves from the base used to measure such deviations. The resultant is known as the ratio of variation.

To calculate this ratio of variation we take the deviation of the relative items from the mean at each date and divide it by the corresponding deviation of the subject. The quotients so obtained would then be added and divided by the number of quotients.

This calculation is satisfactory when the movements are regular, but a better method is to plot the Index Figures of the data (columns 1 and 4 of Table Y) using the vertical scale for the Index of the subject and the horizontal for that of the Relative, as shown in the accompanying diagram known as the Galton graph. We thus obtain a number of points, some widely scattered. The next step is to draw a line most nearly approximating the general trend of the points plotted. This is generally done by finding a line running in the direction, as nearly as can be ascertained by the eye, and having the same number of points on either side and, as far as possible, equidistant from it. If perfect correlation be present the line obtained by joining the points plotted will be perfectly straight, or, if there be a lag, a well defined and regular curve.

If the line points downward to the left then the correlation is direct, and *vice versa;* while if no well-defined tendency is exhibited no correlation is present.

In using this Galton graph we must remember that if the relative changes by, say, 5 per cent., and the

subject changes to a like extent, then the ratio of variation is obviously unity. In such a case the points will be found arranged on either side of a line of 45° slope, and such line represents a line of *equal proportional variation.* When the Relative shows a tendency to change less than the Subject, the line will be at a less angle than 45° to the vertical. This line is known as the *regression line.* The nearer this regression line approaches the vertical the slighter the degree of correlation. The larger the number of points plotted the more reliable the results will be.

Example.

Plot a Galton graph from the following table and show the Ratio of Variation between Bank Clearings and Immigrants for eight years.

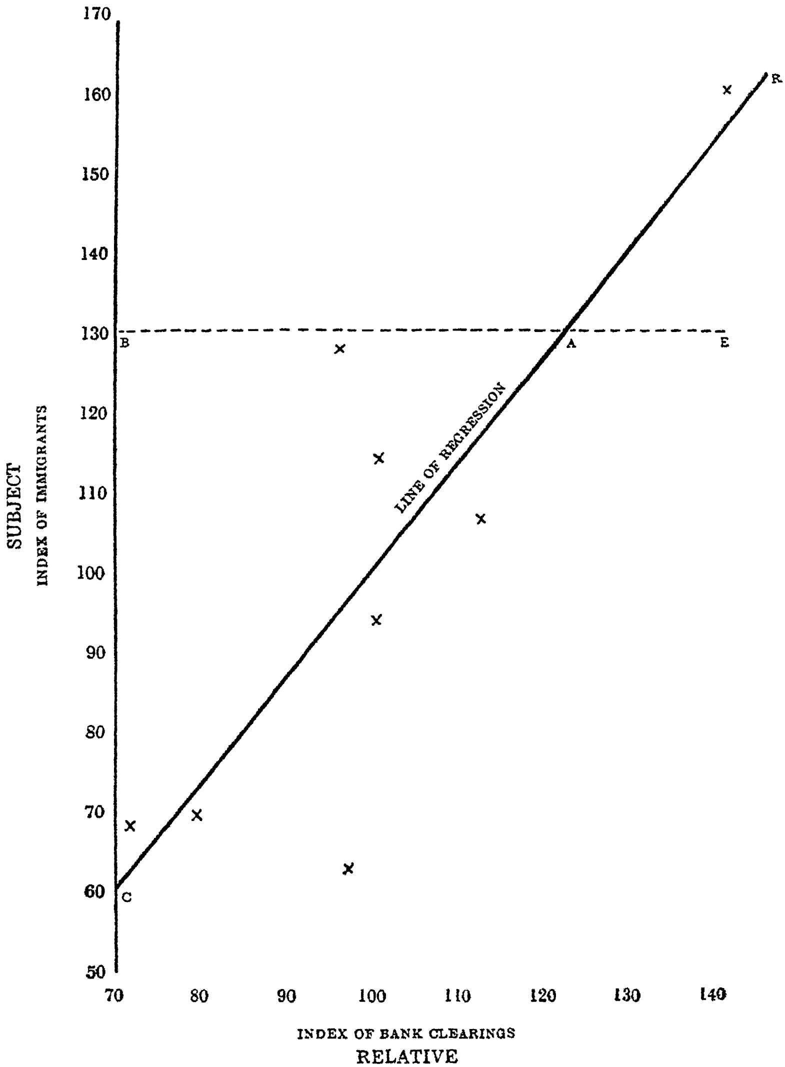
Year.	Immigrants in Tens of Thousands.	Bank Clearings in Millions of £'s.
1	79	49
2	52	40
3	33	25
4	55	35
5	46	35
6	62	34
7	31	34
8	34	28
	Average = 49	35

Using the average as the base in each case we get the following Indices:

Year.	Index of Immigrants.	Index of Bank Clearings.
1	161	140
2	106	114
3	67	71
4	112	100
5	94	100
6	126	97
7	63	97
8	69	80

These are then plotted and the Line of Regression drawn (CR).

Diagram No. 25.
GALTON GRAPH.



To find the ratio of variation draw a horizontal line at any point such as BE, cutting the Line of Regression at A. The ratio of the average variation of the Relative to the average variation of the subject is represented by AB/BC or the tangent of the angle ACB.

In the graph AB = 52 and BC = 70.

$$\therefore \text{The ratio of variation} = \frac{52}{70} = 0{\cdot}74 \text{ approx.}$$

Thus for every change of 1 per cent. in the subject there is a tendency for the Relative to change ·74 per cent. The complement of this fraction, viz., 0·26, is called the *ratio of regression.*

SYNOPSIS TO CHAPTER XII.

THE CENSUS AS AN AID TO SCIENTIFIC BUSINESS.

1.—Information obtainable from the Census.

2.—Value of Census Statistics to Business Men.

3.—Density of Population and Market Possibilities.

4.—Per Capita Consumption of Commodities.

5.—Estimating the Quality of Markets.

6.—Occupational Lists and Labour Supply.

7.—The Localisation of Industries.

8.—Value of Census Reports to Producers and Manufacturers.

CHAPTER XII.

THE CENSUS AS AN AID TO SCIENTIFIC BUSINESS.

The Census taken by the Registrar-General of this country is primarily intended for the use of the Government, but it is of very great interest to economists and students of sociology, and affords much material for study by statisticans. The preliminary report of the Registrar-General shows the distribution of the population, as well as the increases and decreases which have taken place in different areas. Business men, however, do not, as a general rule, recognise that in this report, as well as in the others which have yet to be compiled and published, there is information of a very important nature; information which, if studied carefully and intelligently, may solve some of the problems of business which confront them.

§ 1.—Information obtainable from the Census.

The Census is usually taken every ten years, but it has been proposed to count the population every fifth year. A form is distributed to each household

calling for information relating to every person therein at a particular date, under the following headings :—

(1) Names of the persons comprising the Household.
(2) Relationship to the head of the house (*e.g.*, Wife, Son, Visitor, Boarder, Servant, etc.).
(3) Age in years and months.
(4) Sex.
(5) Whether married, single, widowed or divorced. In the case of persons under 15 years of age, whether or not parents are alive.
(6) Birthplace and nationality.
(7) Personal Occupation.
(8) Employment.
(9) Place of Work.
(10) Number and age of all living children and step-children under 16 years of age, whether residing in the house or elsewhere.

From this information it is possible to ascertain not only reliable figures as to the number of inhabitants in this country, but also how they are distributed, and the sorts and conditions of people who form the population of the various areas into which the country is divided. A study of this information enables the true condition, both social and economic, of the people to be ascertained. These data enable not only the national but also local government to be adjusted to the peculiarities shown. The details as to the age of the inhabitants permits of estimates as to the amount of money which will be required to provide for old age pensions, and for the provision of educational facilities, to be made with a reasonable degree of accuracy; while comparisons of the number of deaths per thousand

of the population in each area will supply clues for medical investigations into health and hygienic conditions. The distribution of the population will show the centres most suitable for the distribution of foodstuffs essential to the welfare of the inhabitants in times of transport difficulties, and will also enable the demand for the necessities of life to be estimated accurately; while both national and local services of public utility can be provided according to the varying needs of the different districts.

§ 2.—Value of Census Statistics to Business Men.

The trade of this country depends more than anything else upon one factor, the *consumer*, and the manufacturer, wholesaler, or retailer, needs to study his customers very closely indeed if he is to build a big business and create a valuable goodwill. If the goods supplied do not meet with the approval of the consumer they will not sell in large quantities, however cheaply they may be produced, while it is obviously useless and a waste of money to endeavour to sell articles of luxury in an area the inhabitants of which consist largely of unskilled workers whose incomes are relatively small. The Census Reports supply a wealth of information about the consumers and where they are located, and thus will well repay a business man who carefully studies their contents, if only from this point of view alone. The object of the writer is to point out some lines for investigation for those interested to follow. Some of these will be found to apply to certain trades or industries only, but every trader or business man can obtain some useful information if he will but study the reports as they are issued.

§ 3.—Density of Population and Market Possibilities.

In the first place, the reports will show the distribution of the people, and since the areas of the various districts are also included, it is possible, with but little calculation, to ascertain the density of the population in any given area. To the producer of goods, and to the large distributing houses, this will afford invaluable information as to the likely areas in which to develop a market, for, *all other things being equal*, the best district upon which to concentrate the efforts will be that one which has the greatest density of population. An example will perhaps be of interest. Supposing it is desired to open a branch establishment, say, in Yorkshire, a county which contains several large towns and cities. The preliminary Report of the 1931 Census showed that the two predominant boroughs were Sheffield, with a population of 490,724, and Leeds, with 458,320 inhabitants. The former town, however, has an area of 24,930 acres as against 28,090 acres in Leeds. It may, of course, happen that suburbs which are separately shown, being independent boroughs, may alter the ratio when combined with the figures for the main areas, but, *again all other things being equal*, Sheffield would appear to be the better centre, for not only is the population greater, but its area is smaller, thus reducing the scope and also the expenses of the delivery service, since the density of the population is so much greater. The question of suburbs is, however, an important one in the case of a large town, for although these are actually outside the borough or city boundaries, yet their inhabitants may be engaged in the town itself, and are thus prospective customers. This fact needs no demonstration when we consider the case of the City of London, with a

resident population in 1921 of only 13,706. The geography of the district therefore needs to be studied, and the area and population of the surrounding suburbs taken into account, while if the object is to establish a distributing centre, the proximity of other large towns or populous areas must not be overlooked if the greatest amount of business is to be done economically. The information as to the distribution and density of the population may be said to show the "Quantitative" market.

§ 4.—Per Capita Consumption of Commodities.

Where an established business has already widespread sales, the total sales for various areas can be ascertained with but little trouble, and the *per capita* consumption for the different areas calculated by dividing the consumption of the district by the number of inhabitants therein as shown by the Census Reports. A comparison of the *per capita* consumption of neighbouring districts will show whether there is any marked divergence; and if such is found to exist, then steps can be taken to level up the consumption by investigating the causes and planning an intensive selling and advertising campaign. If the particular commodity handled appeals only to certain sections of the community, then the *per capita* consumption of that section can be ascertained and compared in like manner; for instance, if the goods are intended for the ladies only, the figures relating to the female population would be used, while when the analysed returns showing the distribution of the inhabitants by ages are available, other data may be obtained. For instance, the number of infants will be available, and also the areas where the infantile population is dense

will be shown, and these figures can be used if the product is one intended for the consumption of children, or deducted from the total population where the goods are for adults only. Similarly, by carrying the analysis of the sales further, it will be possible to ascertain whether the *per capita* consumption varies at different times of the year, and this information can be used with a view to discovering whether the article is one suitable for use all the year round, and taking steps to increase the sales in the seasons when consumption decreases, or, alternatively, of locating the limits of the seasonal demand in the various areas.

§ 5.—Estimating the Quality of Markets.

Detailed reports by the Registrar-General will show the occupations of the population, and how persons following the various professions, trades or businesses are distributed over the different towns and districts. From such information it will be possible to estimate the "quality" of the markets available in the different areas, for if the population in one centre, though dense, consists of relatively poorly paid workers, it may not be so advantageous or fruitful in results to a business as another district with fewer inhabitants, each of whom earns higher wages than those ruling in the other area. To obtain such information it will be necessary to supplement the information obtained from the Census Reports with the statistics of wages and hours of employment issued from time to time by the Ministry of Labour. From these figures a fairly accurate estimate can be made of the income of the working-class areas, and thus the purchasing power of the community in such areas can be gauged to some extent. Apart from this, however, the regional lists of occupations will supply clues as to the type of commodity likely to

be demanded within any given area or areas, and thus arrangements be made to meet the possible demand which may arise, or, alternatively, steps taken to interest the inhabitants in the particular commodity, and so create a demand.

In the cases of those houses producing or handling a product which appeals only to a certain class or classes of the community, the localities in which such people are densely congregated will be shown in the regional lists, and the necessary steps to develop any new fields so revealed can then be taken. In each of these cases energies will be concentrated upon the most likely districts to produce results, thus permitting advertising allocations to be adjusted and used to the best possible advantage. Wasteful expenditure on this very important item will thus be avoided.

Where a business has a purely local demand for a particular commodity or type of commodity, the principal of such a house can consider the advisability of extending his operations to areas having similar characteristics, and these can be found by studying the distribution of the population by occupations, added to a study of any geographic conditions which may have an effect upon the demand. This geographic factor is naturally an important one in many cases, for a commodity which suits a certain class, in, say, the South of England, may be totally unsuited for consumption by a similar class in the North of Scotland.

§ 6.—Occupational Lists and Labour Supply.

The occupational lists will also show the apparent supply of labour, both skilled and unskilled, available for different industries in the various districts, and this is an important factor to be considered when deciding

the establishment and location of a new works. Apart from climatic conditions it would, for example, be foolish to commence a cotton-spinning or weaving mill in Kent, for the supply of skilled labour would be non-existent, and thus, unless such labour was very mobile indeed, or special inducements were offered, the project would probably fail from this cause alone. It must also be remembered that labour is not likely to be attracted to an undertaking unless there are reasonable facilities for changing the scene of work in the same district, should the necessity to do so arise, otherwise, should the particular business fail, the position of the worker is much worse than when other works of a similar nature are in the near neighbourhood. Similarly, the report, showing the distribution of the population by age and sex, will be useful for studying the problem of future labour supply. It will be found that in those industrial areas where the rate of wages is good, and continuity of employment is the rule, the great tendency is for the majority of the children to pass into the industry. Such is not the case, however, where seasonal or other fluctuations cause the periods of unemployment to be considerable. The future labour supply is an important factor, and in establishing a new industry or works one would naturally avoid areas where the ratio of children to the total population is low, or where the fertility of marriage is small, in favour of those districts showing a large proportion of children or a high ratio of productivity. Another factor which has to be considered when studying the labour supply is the number of employers, and this information is also obtainable from the reports. If it be found that the number of employers is increasing, while the labour supply is either stationary or decreasing and the future supply

of labour, therefore, is likely to diminish, competition for the available supply of workers will ensue, high wages will result, and the business may not therefore prove to be as remunerative as would appear upon the surface. Anyone, for example, thinking of establishing a new private educational institution, should carefully study points such as these, unless, of course, the personal connections are of such a nature as to insure immediate and future success.

The third factor which is of great importance when establishing a new business in the home trade is the likely demand, and this also may be gauged to some extent by a study of the Census Tables, as already pointed out when considering the question of possible markets, and therefore that business which is established in a good centre for labour supply, without too many competitors, and near a good market for its product, has a better chance of success than one not so favourably situated.

§ 7.—The Localisation of Industries.

Still another factor needs consideration, and that is the proximity to supplies of raw material, for obviously these should be cheaper to purchase when produced in the neighbourhood than when they have to be transported long distances. Here, again, the occupational lists should supply quite a lot of useful information, for it will be possible to ascertain where goods or material are produced by studying the distribution of the workers, for, when a certain class of worker is found to be in a certain area the location of the industry may be assumed. Information regarding the localisation of industries is always of importance to a new concern, for, in some cases, there may be climatic or physical reasons for such localisation,

reasons which will materially assist in the successful establishment and development of a business, while in others it shows possible fruitful fields for commencing business free from the stress of competition. Any change in the localisation of industries needs to be noted, and the reasons for such changes ascertained. To the purchaser of goods it may show fresh sources of supply of raw or other materials as yet untouched by him. The producer needs the information, since his supply of labour may be affected materially by migrations to other areas developing the industry, and able, owing to natural or other advantages, to produce more cheaply or pay better wages.

The report as to the distribution of the people according to their respective occupations is also of great importance to the large producer, for it is in this return that he will be able to find the number of distributive businesses interested in his product. Unless his policy is one of appointing sole district agencies, every retailer in his line is a possible purchaser and distributor of his goods, and he can thus ensure that all are properly approached in order to secure their active co-operation for their mutual profit.

Advertising men may likewise reap much information from the distribution of the population, the areas of greatest density showing the most likely field for general advertising, the occupational lists giving details of areas suitable for advertising commodities of interest to special classes or industries, and generally should assist in suitable forms of advertising being adopted, for the type of advertisement which attracts replies and business from an agricultural area may bring no result at all from a mining or manufacturing area. Luxury articles need advertising in appropriate localities, furniture and other household articles in

areas where the number of young people of marriageable age is considerable, and articles for children in districts where they are most numerous.

To those concerns engaged in transport of any kind the reports should be very helpful, for the distribution of the population and the location and magnitude of industries will show fields capable of development, and investigation will reveal opportunities for extension of existing facilities, altering the present services to meet changed conditions, and so benefit both the public and the business.

§ 8.—Value of Census Reports to Producers and Manufacturers.

The producer of articles of staple consumption can see from the reports where to concentrate his efforts in order to develop his markets, and also the possible number of consumers he may hope to reach. The manufacturer of a household commodity by studying the distributions by families will obtain information to guide him as to likely sources of demand, and can plan his selling campaign to reach the maximum number with the minimum of expenditure. The sales manager by studying the figures given can possibly account for decreases which have taken place in his sales in certain areas by the decrease shown in the population therein, and can thus relax efforts to maintain sales in a district which can only be expected to purchase less. He can thus use his efforts, and expend money more usefully, in giving more attention to areas which show considerable increases. A study of the preliminary report shows, for instance. that the population of the county of Cornwall decreased by no less than 7,539 people, the decline in Penzance alone accounting for 1,382 of this total. It will be found that in this particular town the population has fallen

10 per cent. on the figures at the previous Census. All other things being equal, therefore, sales in Cornwall, and Penzance in particular, must be expected to show a decline unless the remaining population increase their consumption *per capita*, but, knowing the true circumstance of the cases, there would be no need to worry over such fall. Similarly, we find that Dover shows a decrease of 3,660, while the neighbouring town of Folkestone shows an increase of 4,069, consequently efforts should be directed to increasing sales in the latter area rather than try to maintain the same turnover in Dover. Other areas will be found to show considerable increases, and inquiry would therefore be made to see if the sales had increased in like ratio, for, if they have not, steps should be taken to develop the latent market which is thus shown to exist, and travellers' routes and advertising allocations adjusted to the changed conditions thus revealed; for growing areas need "nursing," and declining districts while not neglected should not receive the same attention as heretofore, provided that the *per capita* consumption of the population does not show a decrease at the same time as the number of inhabitants. Demand in certain areas for certain commodities may show a decline owing to a change in the sex composition of the population, for if the ratio which the two sexes bear to one another changes, the character of demand is likely to change also. In Devonshire, for example, the population increased by 9,785, but the male element shows a decrease of 787, while the fair sex show an increase of no less than 10,572, so that the demand for feminine commodities should show a marked increase as a consequence, while the male demand might justifiably be expected to decline slightly. The other factors as to age and industry must, however, be considered at the same time.

Even the local trader can study his own area to see what changes have taken place in the distribution and occupations of the inhabitants, for these may affect his trade materially. An increase in the population should increase demand; a change in the occupational pursuits may change the nature of the goods in demand. If infants are numerous he can lay himself out to supply their multitudinous needs; if young people of marriageable age predominate, he can provide for their requirements; if people of mature years are present in considerable numbers, attention can be concentrated on those commodities which appeal to old age. If the area is a growing one, he must take steps to see that his demand grows at least in the same ratio as the population. If it does not, he is being left behind in the business race.

In studying the preliminary report it must be remembered that the later and more detailed returns may alter the aspect of the first presented data in some cases. The forced postponement of the 1921 Census to a Sunday, undoubtedly affected the distribution of the population to some degree, for holiday makers were recorded at the holiday resorts instead of in their own towns, while a Sunday in summer results in many people being away from home on week-end or other short visits. The occupational lists will, however, be prepared on the basis of the place of employment, and will thus correct to some degree the figures for seaside towns and other holiday resorts, and therefore the figures for these must only be tentatively accepted.

Numerous other uses for the Census figures will doubtless reveal themselves on intelligent inquiry, but enough has been said to indicate their great value to enterprising business men.

SYNOPSIS TO CHAPTER XIII.

BUSINESS RESEARCH.

§ 1.—Profit Earning Possibilities of Research Work.

2.—Co-ordination of Producing and Selling Organisations.

3.—General Functions and Objects of Business Research.

4.—Qualifications of a Business Research Worker.

5.—Important Statistical Comparisons.

6.—The Organisation of Production.

7.—The Problems of Selling.

8.—Studying Competitors.

9.—Other Factors to be Studied.

CHAPTER XIII.

BUSINESS RESEARCH.

During the last decade much progress has been made in the methods of conducting business progress, which the war retarded in some directions and developed in others, but this progress has been chiefly on the side of production rather than that of administration, and consequently there is to-day a tendency for the attention given to trade and commerce to be of a one-sided nature, instead of developing the whole in its proper proportions.

§ 1.—Profit Earning Possibilities of Research Work.

Manufacturers and producers have long since recognised the necessity for improved methods in the factory and workshop, and many useful inventions have been developed by the desire to produce better goods more quickly and at a lower cost. The war saw a careful study made of commodities, and the provision of substitutes for those which were either unobtainable or scarce, and to-day as a consequence many large industries maintain Research and Experimental Departments controlled by trained and expert experimenters and research workers. The dyeing, chemical,

and metallurgic industries are good examples of expansion resulting from such a policy. The electrical equipment manufacturers spend large sums of money in improving the machinery and equipment which they produce, and the public benefit by better commodities and greater reliability in those on the market. This has been particularly noticeable in the development of the electric lamp. The internal combustion engine of to-day is a reliable unit in the transport world as a result of constant research work, while air transport is an established factor as a consequence of the experimental work of aviation pioneers such as the Brothers Wright, Cody, Count Zeppelin, and others. The experiments and researches of Marconi, Edison, and other scientists have made the world richer in means of communication, etc., and provided ways of utilising capital profitably, as well as opening out great possibilities in other directions. The efficiency of the typewriter, calculating machines, cash registers, and other mechanical aids to business is due to the constant efforts to overcome difficulties and increase the service to be obtained from them. All these are sound business propositions resulting in the profitable employment of capital, and this result is due to the untiring and never ceasing efforts of the research workers. Industrial research is to-day an established factor in manufacturing, its utility is recognised, and its profit-earning possibilities demonstrated to the world.

§ 2.—Co-ordination of Producing and Selling Organisations.

Many of the greatest inventions of the age are due to the desire to supply such wants as have been discovered to exist. In other words, the research and

experimental work has been undertaken as a consequence of an inquiry into the needs of the public, and the most successful business is one which supplies the requirements of the general public, but it must be borne in mind that the mere production of an article is of no use unless the public know of its existence, and where and how they may obtain it. To place an article upon the market, even when its utility is beyond question, calls for the exercise of much skill and the expenditure of considerable sums of money. It is therefore becoming increasingly recognised that the selling side of the business needs as careful organisation as does the producing side. The present tendency, however, is to treat each of these two departments as a separate entity, whereas they are interlocked and dependent upon each other to such a degree that both must be considered at one and the same time, and the effects of changes in either upon the other studied. If the selling side is better organised than the producing side of the business, the sales will tend to exceed the productive capacity, and the reputation and results of the business suffer through delay in deliveries; while if too much attention is given to the production, overproduction may result and heavy stocks accumulate. The former is certainly the lesser of the two evils, for the latter must ultimately lead to financial stringency, and frequently to the appointment of a receiver.

The production and the sale of the commodity handled, though of primary importance, are not, however, the only factors in a business which need to be considered. There is also the financial aspect to be remembered, and unless this has also received due attention the other factors may be materially handicapped. It is therefore of great importance that the whole of the factors be co-ordinated in order that

each may receive its due consideration, as by so doing the welfare of the business as a whole will be ensured. The financial resources will then be utilised to the full, production and sales will receive their full quota of attention, and the various interests will be reconciled, thus ensuring the smooth working of the business machine, the ability to finance all operations adequately, the provision for contingencies, the meeting of all competition from other manufacturers, the safeguarding against loss, the elimination of waste, the utilisation of the plant to its full capacity, and the earning of profits.

Many manufacturers are content if their sales equal their production, and many traders satisfied when their turnover shows an increase, but possibly neither is making as much profit as would be possible were the business administered upon a scientific footing. It is a very difficult matter for the business man to obtain a detached view of his own business. To a manufacturer the production of goods tends to outweigh other factors, while the trader is more concerned with the selling side. Moreover, it is impossible for the busy head of affairs to spare the time necessary to study the effects of various changes, or perhaps for altered conditions in advance. If Industrial Research has yielded so much benefit to the commercial community, as well as to those wise enough to recognise the advantages to be reaped therefrom, why should not a Research Department for the business as a whole be beneficial ? Every phenomenon met with in business has a cause, and therefore every change which takes place should be analysed in order to find the cause or causes responsible therefor, so that when similar conditions are again likely to come into operation the results can be foreseen and provided for. In other words Statistical Methods should be used.

§ 3.—General Functions and Objects of Business Research.

Such a Business Research Department would not confine itself to the conditions ruling in the particular business houses instituting it, but would also study the conditions ruling in the industry, and in the commercial and financial world generally. The causes of all changes should be carefully recorded, and their effects noted, so that tendencies can be forecast and their results either taken advantage of or guarded against. If this were done losses would be minimised, profits increased, capital used economically and advantageously, and the business as a whole, and its proprietor in particular, benefit. All this work is essentially of a scientific statistical nature, and if undertaken by trained observers would leave the proprietor or proprietors of a business, or the management of a company, free to direct the general policy of the business instead of being worried with the work of maintaining a close watch on market tendencies and movements.

The Business Research Department should have no executive function, but should confine its attention to carefully studying all the factors relating to the business, the commodity or commodities handled, the methods of production and sale, etc. It would thus ensure that all departments receive due attention, and would hold the balance between the varying interests involved, and so prevent a one-sided organisation. While not an actual profit-earning department, its one aim and object will be to increase the profit-earning capacity of the business by pointing out factors which should be taken advantage of and those which will need to be guarded against. If an Industrial Research or Experimental Department were already in existence

it would in no way interfere with its operations, for the objects of the two are quite distinct. The Industrial Research Department will, as now, concentrate its efforts on improving the methods of production, or in experimenting with new processes or products. The Business Research Department studies the results achieved by the Industrial Research Department to see whether they are commercial propositions and, if so, what effects the manufacture of such products will have upon the business as a whole, and the demand which can be expected or created. It would thus act as a connecting link between the producing, selling, and financial sides of the business, recording, studying, and reporting upon the results which can be anticipated or achieved by each, thus eliminating friction and overlapping of interests and leading to greater efficiency all round.

§ 4.—Qualifications of a Business Research Worker.

Much of the work of such a department would be of a necessity of a statistical nature, and therefore the person selected to take charge of it must have an extensive knowledge of the principles of this important subject. It is, however, essential that the records compiled should be of such a nature that they will yield the results required of them, and not that they will be regarded merely as statistics. The hackneyed use of statistics has brought statistical methods into some disrepute, and this has been very largely due to the fact that the convenience of the persons using the data has not been studied, nor the uses to which they could be put given consideration. The Business Research worker therefore needs to be a business man

as well as a trained observer. His knowledge of business methods and requirements must be wide. He must see his own business house in perspective, and therefore know what to record and what to discard. The principles of banking and accounts must have been studied by him, while the elements of economics, particularly those relating to the operations of the economic laws of supply and demand, and of increasing and diminishing returns must be known. As he will be responsible for assigning due weight to each department he must be tactful in his dealings with the executive officials, confident and purposeful, rapid in his methods and in the presentment of his reports, and capable of demonstrating clearly and succinctly the reasons for the recommendations he makes. His reports must be models of expression, simple in their language, and easily followed by those not so well trained in methods of observation as he himself. His recommendations must produce results, and tabulated and other statements must be prepared with this object in view, rather than for demonstrating his ability as a statistician. He should be essentially a consultant rather than an operator, an advisor than an executive officer.

§ 5.—Important Statistical Comparisons.

Upon the establishment of such a Business Research Department, it should make an exhaustive study of the existing records of the business, as well as the past results. This will enable it to ascertain the reliability of the data at its command, and of the progress or otherwise of the business, and the efficiency of the methods used can thus be gauged. The results achieved by competitors should, whenever possible,

be studied and compared with those of the "house," in order to ascertain whether or not the business is showing results which are equal to the average ruling in the industry. Such comparisons are extremely valuable, and form a useful basis for more detailed investigation; thus all results published by others engaged in the same industry or trade, as well as official and semi-official returns, must be assiduously collected, analysed, recorded, and compared. The financial position and methods of the house will be carefully studied to see if its resources are adequate for the volume of business transacted, or whether there is over capitalisation—an evil which must be guarded against as carefully as that of having too little capital. Attention must also be directed to the policy adopted in regard to the provisions to be made in order to finance an expanding turnover, or an increase in the cost of labour or material, for either of these factors will cause additional working capital to be required. The assets of the business should be examined to see if they are liquid, and therefore capable of easy conversion into cash in case of an emergency. The debtor's accounts need to be considered in order to ascertain whether the ratio of bad debts to turnover can be reduced by adopting other methods of inquiry into the financial status of the customers, or an alteration in the methods of extending credit, and whether the terms of credit can be reduced or collections hastened. Inquiry will also be directed to the costs of collection and the methods adopted, and whether these can be reduced or made more efficient.

§ 6.—The Organisation of Production.

The production of purchasing departments of the business will be examined to see if the methods in

operation are as economical as is possible. The rapidity of production will be inquired into, and divergencies in the time and/or the cost of similar processes or units noted and inquired into. The elasticity of production should also receive attention, to ascertain whether in the event of heavy demands arising they can be met without undue delay. The quantities of stock consumed over regular periods will be periodically compared with the purchases of raw material to see that there is no delay in production owing to lack of raw material, or that large stocks are accumulating, thus locking up working capital and preventing its being used more remuneratively in some other direction. The sources of the supply of material should be gone into with a view to discovering whether any other supplies are available should the usual supplies fail or diminish, and also whether efficient substitutes can be utilised when there is a shortage of the commodities used, or a considerable increase takes place in the price. Records should be made of the fluctuations in the price of raw materials to discover if these are regular in their sequence, and if any advantage can be derived from purchasing from any particular source at any one period. The world's production of the particular raw materials used should be studied to ascertain if these are increasing or decreasing in volume, and what effects such change in production will have upon the business or the policy to be pursued. Labour conditions and the supply of workers available must necessarily receive attention, for any impending changes in wages will have a marked result upon the cost of production. If the supply of workers in any trade is in excess of the demand, wages will tend to fall, in spite of the efforts of trade unions to maintain them, while if unemployment

is small there is likely to be competition for the labour available and wages will tend to increase. Inquiry should also be directed to the changes which take place in the labour force of the organisation, for, if there are constant changes in the staff, efficiency and output suffer till the new workers get accustomed to the conditions ruling.

§ 7.—The Problems of Selling.

Many problems of equally great importance will come up for solution when the marketing or selling side of the business comes under review by the Research Department. The existing methods of marketing the goods will be studied and tested, other ideas will be examined and tried in order to find that which is the most efficient and economical, taking care, however, to see that too much economy is not practised, for this may have a detrimental effect upon the turnover of the business. The past records will be tabulated, and examined to discover whether the turnover is progressive, care being taken to compare both quantities and value, for if prices have changed, the value of the turnover may have been affected, although the quantities sold may show a movement in the opposite direction. The number of customers supplied and the average amount and quantity of their purchases will also be scrutinised, for here again valuable information as to the progress or otherwise of the business will be available. Sales records will be analysed to find whether the demand for the products is constant or fluctuating, and if the latter whether the fluctuations are of a regular or irregular nature and the reasons therefor. If the demand shows

an increase at regularly recurring intervals, arrangements can be made to adjust production accordingly, and investigations made as to the possibility of regularising the demand, or of utilising the plant on the production of other commodities during the slack periods. If the demand is found to be increasing, it should be seen that satisfactory arrangements have been made to meet the rising consumption. If demand be falling, the reasons should be inquired into, and attention directed to the possibility of stimulating sales, or of altering the product to meet changing tastes on the part of the consumers. The Research Department will study the consumer closely with a view to discovering whether the demand is limited to a particular class or classes of the community, or is of a general nature. If the former proves to be the case, tests should be made as to the possibility of interesting other classes of the community. The demand may be of a local nature, or confined to certain districts, and here again it will be necessary to ascertain the reason for this phenomena, and efforts directed to creating a demand in other localities. Statistics relating to the consumption *per capita* in the varying districts will at once indicate those areas to which special attention should be devoted. In cases where the sales are confined to the home trade the possibilities of cultivating foreign markets should be investigated, as by widening the area over which sales take place a steady turnover can be partly assured, seeing that, except in extreme cases, factors in operation are not likely to affect conditions in all countries at the same time or in the same degree. Some districts or countries may be found to have a seasonal demand, and special arrangements should then be made to supply such

areas at the time the demand occurs. The routes of the travellers should be inquired into to see that every centre is receiving attention, and that the energies of the representatives are devoted to increasing the number of customers, as well as taking orders from those already on the books. The local tastes must be a subject for review, as it often happens that the desires of consumers are not met, and sales diminish as a consequence. The organisation must be studied, and the possibilities of better results being obtained from the establishment of branch establishments, or sole agencies exhausted. It may be found that increased sales may lead to decreased costs of production, owing to the plant being more fully occupied, and tests should be made as to whether a decrease in the selling price will lead to an increased demand, and whether this increased demand allied to a decrease in the costs of production will be to the financial advantage of the business.

The total supply of the finished goods available from all sources is also a subject for inquiry. If competition be increasing, the effect which this is likely to have upon the selling price must be considered, and the result which this will have upon the business ascertained or estimated. The advertising department will also be reviewed with the object of ascertaining the particular methods applicable to varying conditions and different districts. The allocation of the advertising expenditure will be ascertained, and the results traced wherever possible. It will be possible to find the best media for advertising, and the periods which produce the best results. Tests should be made with various styles of advertisements to ascertain that which attracts the most business.

§ 8.—Studying Competitors.

Not only must the business in itself be studied and analysed, but the industry as a whole considered, for if there is a slump in a particular trade it may be a wasteful process to run an advertising campaign at such a time, but if conditions are known, and the reasons therefor ascertained, the difficulties may be overcome and sales maintained. A study of the competitors, the ground they cover, and their methods of conducting business will often supply a clue to the reasons for a falling turnover, while if the number of competitors increases without the general demand changing to any marked degree, special steps may have to be taken to maintain the volume of trade, and here again the Business Research Department, with its records, should supply invaluable assistance in deciding upon the course to be pursued, or alternatively may be able to suggest some variation in the form of the commodity which will have greater possibilities of profit. Trade and commerce in general must also be included in the survey carried out by this department, for a national slump may be accompanied by a boom in some other country, and hence the poor returns shown by the home trade may be offset by an increase in the export sales. The varying conditions ruling abroad must also be studied in order that advantage may be taken of the most favourable markets, and so enable both turnover and profits to be increased.

§ 9.—Other Factors to be Studied.

Financial and credit conditions must not be overlooked. A stringency in the financial world may result in longer credit being asked for, or in credit facilities having to be withdrawn or modified according to the

conditions ruling, and therefore if losses are to be guarded against, the economic conditions ruling throughout the world must be carefully studied and the results foreseen. If, for example, the world's cotton crop is a poor one, Lancashire will suffer considerably, trade in that part of the country will decline, and the credit policy for that district must be decided upon before the slump actually makes its appearance ; but the Research Department, if properly organised, would have given warning of the impending difficulties and provision would have been made accordingly. If a financial crisis appears imminent in any country, the magnitude of its effects must be forecasted and guarded against if loss is to be prevented. Foreign exchange fluctuations may turn an apparent profit into an actual loss unless they have been carefully studied and the reasons for the movements ascertained.

Enough has been said to demonstrate the many points on which a Research Department could focus its attention in order that business may be conducted on strictly scientific lines. Such a department would supply all the information needed to enable a busy principal to decide the methods to be adopted and the policy to be carried out, and would thus act as an intelligence department to the chief of the business. It may be argued that only large businesses could afford to maintain such a department, and this is in a large measure true, but the various associations of employers could do much to supply general information as to local, district, national, or international conditions of trade, commerce, and finance, and their effects under the particular industry as a whole, leaving each head of smaller businesses to apply the information to his own needs. The Federation of British Industries has established such a department, the operations of

the Commercial Gas Users Association are well known, while the Universities of London and Cambridge have established an Economic Bureau known as the **London and Cambridge Economic Service**, which works in conjunction with the Harvard University Committee on Economic Research. The object of this service is to render to business men expert analyses of industrial conditions, to supply statistics and memoranda similar to the services which the Universities have long rendered in the fields of chemistry and other natural sciences. It issues monthly bulletins of carefully chosen statistics, with Tables and Diagrams, showing clearly the movements of the main elements in the financial and industrial position up to date, with a summary and analysis of the changes, and, so far as is possible, a forecast of the near future; while periodically or occasionally it issues memoranda on special topics, such as the volume and movements of foreign trade, the course of the Foreign Exchanges, stocks of commodities, the condition of the Money Market, the position of freights and shipping. In America there are several organisations giving similar services to this, and apart from Harvard University the principal are :—The Brookmire Economic Service, The American Chamber of Economics, and Babson's Statistical Organisation.

SYNOPSIS TO CHAPTER XIV.

MECHANICAL AIDS TO STATISTICAL WORK.

§ 1.—Calculating Devices.
2.—The Powers Sorting and Tabulating Machinery.

CHAPTER XIV.

MECHANICAL AIDS TO STATISTICAL WORK.

Many Statistical Returns now presented to the public by Government Departments and other Public bodies suffer from the very serious defect that the results are only available after the lapse of such a period of time that the statistics cannot be used to the best advantage, even if they can be used at all. In other words, while they enable a survey of past movements to be conducted, they cannot be applied to the conditions immediately following the period to which they relate. In commercial statistical work it is essential that the collection, recording, investigating and presentation should be so speedy that the results obtained can be applied without undue delay to the conditions then present; while in research work the facilities for classifying, grading and tabulating should be such that the risk of mistakes is eliminated or minimised, and rapidity of working assured. These two factors can only be achieved when mechanical assistance is available.

§ 1.—Calculating Devices.

There are numerous devices now available to assist the Statistician in efforts to obtain both accuracy and speed, and many of these are well known and have been tested over long periods. When long lists are required for tabulation, or for the purposes of summation, the Burroughs Adding and Listing Machines cannot be excelled; while when it is desired to work computations

involving addition, subtraction, multiplication, division, the raising of powers of a number, or the extraction of a root, machines of the type of the Monroe Calculating Machine, the Comptometer, the Madas or the Millionaire supply the need, and ease the labour involved. Most of these machines, however, suffer from the disadvantage that the figures used are not recorded except on the dials of the machine, and so cannot be afterwards checked to ensure accuracy; thus, unless great care is exercised at the outset of the operation, mistakes may creep in and the result may be affected to an unknown degree. The rapidity of operation, however, is such that provided reasonable care be taken by the operator, the advantage to be obtained from the speed of computation outweighs the disadvantage of possible mistakes, while the size of the operations which can be carried out on such machines ensures a greater degree of accuracy, when dealing with large numbers, than can be obtained by the use of the Slide Rule, and results can be obtained much more quickly than is possible by the use of logarithms.

§ 2.—The Powers Sorting and Tabulating Machinery.

Where large quantities of detailed data are being handled, and particularly where Research work calls for numerous groupings and classifications entailing considerable sorting and resorting of the data, as well as tabulating the results under the various headings, the problem of speedy handling becomes very much more complex. After considerable experience, investigation and experiment, however, the writer has found that the recording, classifying, tabulating and investigating of data can be best carried out on

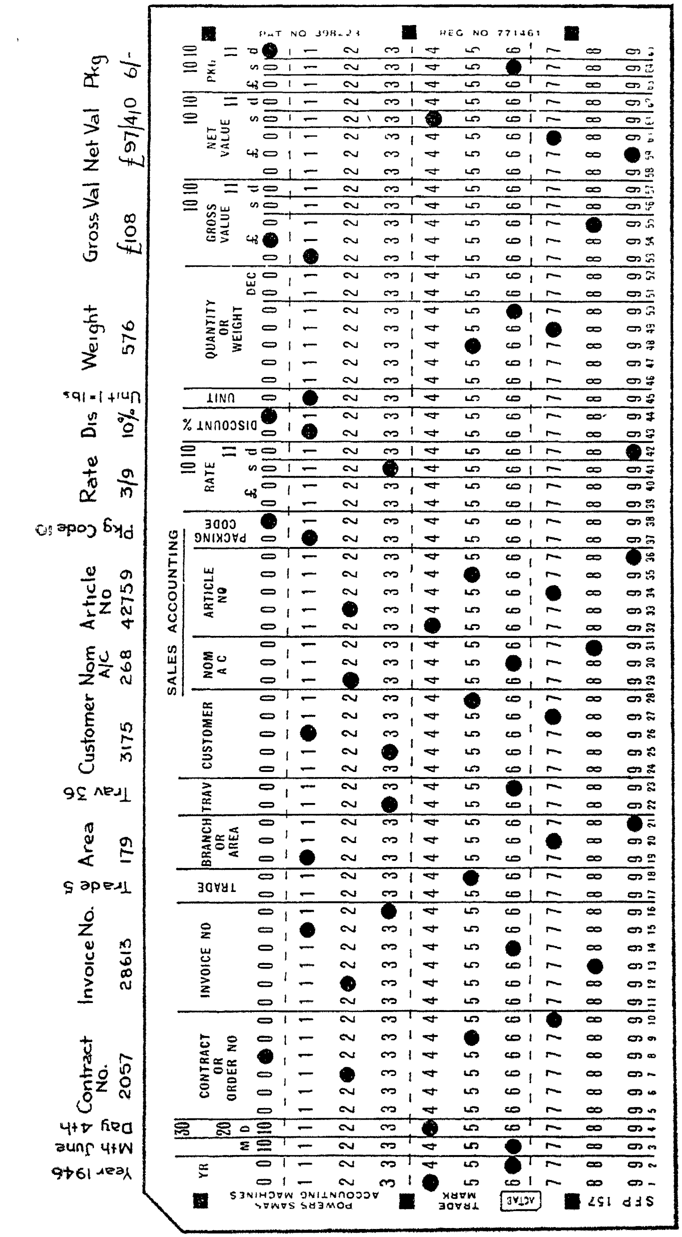


ILLUSTRATION NO. 1.

the **Powers Punching, Sorting and Tabulating Machinery.** The necessary data is punched on a specially designed card (*see Illustration No.* 1) containing sixty-five columns in which punching can take place, each column giving a range of numbers from 0 to 10, though on a large proportion of the columns it is possible to arrange for numbers ranging from 0 to 39 to be punched, while other combinations of numbers can be specially arranged to meet the needs of the user.

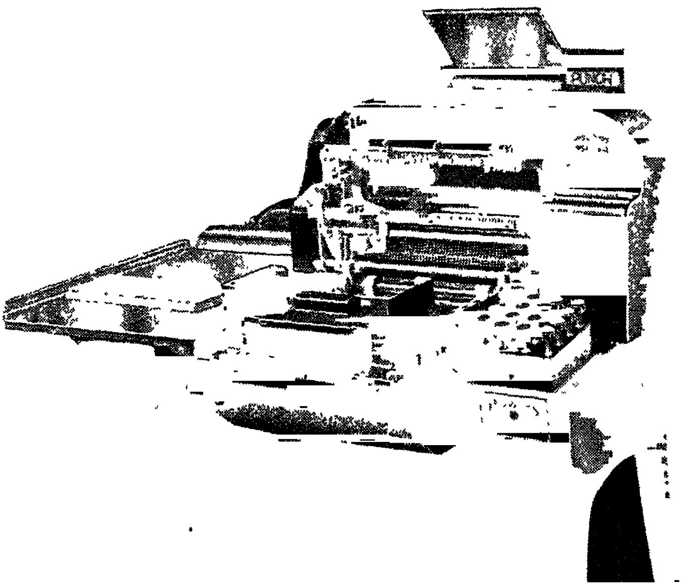


Illustration No. 2.

Although in many classes of work, alphabetical data is now recorded on punched cards, in dealing with statistics, the data is usually coded into a numerical code, each factor being indicated by a previously agreed number. It is only in the punching of the cards that errors can occur, but the system guards against these by incorporating very efficient methods of hand and mechanical verification. Illustration No. 2 shows the Automatic Key Punch which is used for the purpose of

the punching, and an expert operator can easily punch 300 cards per hour from the particulars coded.

The Sorting Machine (*see Illustration No.* 3) is capable of sorting and counting cards at a rate of 24,000 per hour. It operates automatically, and so long as the magazine is kept filled will run continuously with very little attention. The last card passing through will stop the machine. Sorting can

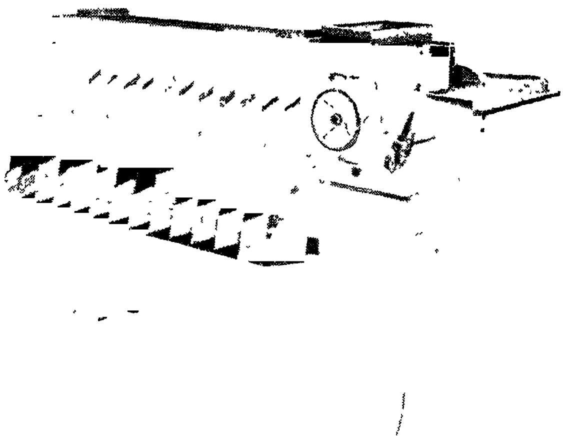


ILLUSTRATION No. 3.

be conducted on any column, and consequently any factor provided for in the code used can be extracted in a minimim of time, and with no possibility of error. The number of examples of each group within the class selected are recorded on the counters, as well as the total data falling in the column on which sorting is being conducted. If the code number for a particular factor is spread over two or more columns, then resorting is, of course, necessary. Supposing for instance, it is required to ascertain the premium income of an insurance company in respect of Ford Cars, the code number for which is 355, and that this information is

recorded in, say, columns 39, 40 and 41 of the card. The whole of the cards relating to premium income first are sorted into ten groups on column 39. Only those cards which fall into the fourth group (*i.e.*, those on which 3

Illustration No. 4.

has been punched) are required for further operations and the remainder are filed. The cards retained are again sorted into ten groups on the punchings in column 40, and those punched with the figure 5 are retained for further sorting, which is then likewise

conducted on column 41. In this last sorting those bearing the figure 5 will be retained, and are those we require to enable us to obtain the information relative to Ford Cars in which we are at the moment interested.

The cards so selected are then passed through the Tabulator (*see Illustration No.* 4), and here the punch holes are converted into figures, and not only printed but are also summated on columns provided for that purpose, and in the example given would supply us with the total premium income for that type of car.

The tabulator will handle some 5,000 cards per hour, and as many as seven different totals can be obtained from the machine. A specimen of the tabulation obtained is shown in Illustration No. 5.

The speed and accuracy of these machines enables the data to be recorded in a minimum of time, and consequently records can be maintained correct and complete to within a few hours of the transaction taking place, while details regarding any factor provided for on the card, subdivided under various headings as may be necessary, can be obtained and tabulated within a short time of the need for the information arising. It would call for but a short expenditure of time and trouble to subdivide the premium income referred to above under the headings of, say, Branches, Months, Class of Policy, Agent, etc. The facility and rapidity with which the data can be classified under any heading provided for in the scheme is one of the greatest advantages of this machinery, and it should prove of enormous benefit to the Commercial Statistician and Research Worker. It may be mentioned that it is the use of this machinery in the work of preparing the reports on the Census that has made it possible to enumerate the population every fifth

MATERIAL COST ANALYSIS BY PART NO.

Week ending ....18th April........

Date d.	Date m.	Req'n. No.	Quantity to be made	Sales Order No.	Bin No.	Part No.	Quantity completed	Unit	Price s. d	Value £.	Value s. d
11	4	19263	18	6283	31256	6320	20	4	7.3	7.	5.0
11	4	19267	27	6302	31256	6320	30	4	7.3	10.	17.6
12	4	19283	45	6315	31256	6320	50	4	7.3	18.	2.6
13	4	19305	18	6329	31256	6320	20	4	7.3	7.	5.0
							120			* 43.	10.0
11	4	19264	500	6287	31258	6321	50	3	3.8	9.	3.4
14	4	19381	500	6372	31258	6321	50	3	3.8	9.	3.4
							100			* 18.	6.8
11	4	19265	6500	6281	29162	6488	750	1	1.3	46.	17.6
11	4	19269	575	6284	29162	6488	50	1	1.3	3.	2.6
12	4	19291	450	6301	29162	6488	40	1	1.3	2.	10.0
13	4	19302	400	6324	29162	6488	36	1	1.3	2.	5.0
15	4	19394	820	6369	29162	6488	78	1	1.3	4.	17.6
							954			* 59.	12.6

ILLUSTRATION No. 5.

year instead of every decade as before, while the Reports are being issued much more rapidly than was previously the case. The use of this plant is not confined to statistical work, for it is in use in Store-keeping, Costing, Traffic Returns, etc., and can be applied to any form of Accounting work in which large masses of data have to be recorded and tabulated.

INDEX

A

ABNORMAL FLUCTUATIONS,
How eliminated for the Purposes of Comparison, 19, 31, 69 75.

ABSOLUTE DISPERSION, 223.

ABSOLUTE ERROR,
Definition of, 53.
Estimate of Biassed, 58.
Estimate of Unbiassed, 56.
In Estimated Average, 59.

ABSOLUTE MEASURE OF SKEWNESS, 236.

ACCOUNTANCY,
Cost, A Branch of Statistics, 7.
Use of Percentages in, 128.
Where it differs from Statistics, 7, 28

ACCUMULATED SALES,
Method of Presenting, 187.
Uses of showing, 189.

ACCURACY, 46.
Degrees of, how shown, 48.
Importance of, 47.
Necessity for showing, 47.
Of Estimates, 47.
Standard of in Estimates, 47.

ADVANTAGES, ARISING FROM THE USE OF THE
Arithmetic Average, 70.
Curve of Normal Frequency, 149.
Decimals, 67.
Graphic Method, 181.
Index Numbers, 242.
Median, 99.
Mode, 107.
Round Numbers, 48.

ADVERTISING,
Records of Results from, 41, 212.

AGGREGATE,
Definition of, 70.

AGGREGATIVE METHOD
As applied to Weighting of Prices, 255.

APPLICATION OF THE GRAPHIC METHOD, 145.

APPROXIMATION, 48.
Effect of on Percentages, 52.
Methods of, 48.
Need for Care when Decimals are being used, 67.

AREAS IN DIAGRAMS,
Care in using for Comparative Purposes, 174.

ARITHMETIC AVERAGE, 70.
Advantages of, 71.
As a Standard of Comparison, 69.
Commercial Uses of, 75, 90
Descriptive, 92.
Disadvantages of, 72.
Effect of Skewness on, 152.
Example of Simple, 70.
In Index Numbers, 245.
Moving Average Standard, 19, 77, 190.
Period over which to be calculated, 75.
Progressive, 81.
Progressive Average Standard applied to Commerce, 90.
Short Method of Computing, 71.
Simple, 70.
Standards of Comparison computed on varying bases, 77.
Uses of in Commerce, 75, 90, 128.
Weighted, 83.

ARITHMETIC MEAN (A.M.), 245.
Compared with Geometric Mean, 250.

ARRANGEMENT OF DIAGRAMS, 182.

ARRAY,
Definition of, 96.
Example of, 142.

ASSESSMENTS,
Effect of on Rating, 18.

AVERAGE,
Deviation, 223.
Error in Estimate of, 59.

AVERAGE CURVES IN DIAGRAMS,
How plotted, 166.

AVERAGES, 69. (See also TYPES.)
Descriptive Arithmetic, 92.
Geometric, 94.
Moving, 19, 77, 190.
Progressive Arithmetic, 81.
Simple Arithmetic, 70.
Uses of, 69, 75, 90, 128.
Weighted Arithmetic, 83.

B

"BANKERS' MAGAZINE" INVESTMENT INDEX NUMBER, 280.

BANK HOLIDAYS,
Effect of, on Output, 18, 31.

BANK OF ENGLAND RETURN,
Example of Tabulated, 124, 125.

BANKS AND THE USE OF STATISTICS, 3, 7.

BAR DIAGRAMS, 162, 212.

BAR DIAGRAMS SHOWING COMPONENTS, 171, 212.

BASE YEAR, CHANGING OF, 258.

BASES OF COMPARISON,
Changing, Effects of, 14.
Changing, Example of, 14.

BIASSED ERROR,
Definition of, 55.
Estimating the, 58.

BLOCK DIAGRAMS, 161.

BOARD OF TRADE,
Indices of Import and Export Prices, 271.
Published Data of, 26.
Returns, Fallacies in Use of, 17.
Wholesale Price Index, 265.

BRANCH RECORDS, 131.

BUDGET,
A Collection of Estimates, 6.
Method of Compilation of, 5.

BUSINESS,
An Application of the Science of Estimates to, 8, 40.
Census as an Aid to Scientific, 301.
Data of can be accurately compiled, 22.
Necessity for Records in, 3.
Statistical Departments in, 9, 28.
Statistics in, 7, 20, 27, 40.

BUSINESS RESEARCH, 315.
Co-ordination of Producing and Selling Organisation, 316.
Factors to be considered in, 327.
Functions of, 319.
Objects of, 319.
Organisation of Production, 322.
Problems of Selling, 324.
Profit earning, possibilities of, 315.
Qualifications for, 320.
Studying Competitors, 327.

C

CALCULATIONS,
Importance of Checking, 130.
Mechanical Aids to, 331.

CAPITA, PER,
Consumption of Commodities, 305, 325.

CAPITAL,
Changes in must be recorded, 16.
Classes of, 16.
How to calculate amount of, for Purposes of Comparison, 16.
Lack of stability in amount of, 19.

CAR MILE,
Cost per, 39.

CENSUS, THE,
As an Aid to Scientific Business, 301.
Density of Population, 304.
Estimating Labour Supply from. 307.
Estimating the Quality of markets from, 306.
Estimating the Quantitative market from, 305.
Information obtainable from, 301.
Localisation of Industries revealed by, 309.
Market Possibilities, 304.
Value to Business Men of Statistics therefrom, 303, 311.

CHAIN BASE, INDEX NUMBERS CONSTRUCTED ON, 252.

CHARGES,
Records of Delivery, 41.

CIRCULAR DIAGRAMS, 174.

CLASSES OF STATISTICAL DATA, 42.

CLASS INTERVAL, 97.

COAL UNIT COSTS, 39

CO-EFFICIENTS,
Correlation, 292.
Dispersion, 223.
Skewness, 236.

COLLECTION OF DATA,
Foundation of Statistical Science, 11.
Methods of. 22.
Rules for, 22.
Uniformity in, 22.

COMMERCE,
Graphic Method applied to, 185.
Progressive Average Standard applied to, 90.
Statistics in, 7, 21, 40.
Use of Arithmetic Average in, 75, 90.
Use of Index Numbers in, 265.
Use of Mode in, 108.
Use of Weighted Average in, 90.

COMMERCIAL DATA,
Difficulties of Compiling, 27.
Homogeneity in, 22.
Tabulation of, 112, 129, 131, 132, 134, 135, 137.

COMPARABILITY OF DATA, 11, 28.

COMPARISON,
Aim of Statistical Science is, 6.
Fallacious, 14, 28, 60.
How to calculate amount of Capital for purposes of, 16.
Methods of, 219.
Standards of, 69.
Standards of, Computed on various bases, 77.
Units suitable for, 31.
Use of Index Numbers for, 241.

COMPARISON OF AREAS IN DIAGRAMS, 174.

COMPARISON OF QUANTITIES AND VALUES, 194.

COMPARISON OF VOLUMES IN DIAGRAMS, 177.

COMPENSATING ERRORS, 54.

COMPILATION OF DATA ON VARYING BASES, 77.

COMPILATION OF STATISTICAL DATA, 11, 22.

COMPLETENESS OF DATA, Necessity for, 11, 16, 28.

COMPLEX TABULATION, 116.
Examples of, 117, 119, 124, 125, 127, 129, 134, 135.

CONSTRUCTION OF DIAGRAMS, Rules for, 182.

CONSUMPTION OF COMMODITIES *PER CAPITA*, 305, 325.

CONTINUOUS SERIES OF DATA, 42.

CONTROL RECORDS, 211.

CORRECTNESS OF DATA, Importance of, 11.

CORRELATION, 288.
Co-efficient of, 292.
Definition of, 288.
Ratio of Variation, 295.

COST, What constitutes, 215.

COST ACCOUNTANCY,
A Branch of Statistics, 7.
Electrical Industry, 39.
Railway Companies, 39.

COST OF LIVING INDEX NUMBER, 275.
Commodities included in, 276.
Imperfections of, 276.
Method of Compilation, 275.
Weights used in compilation of, 276.

COST PER CAR MILE, 39.

COST PER PASSENGER MILE, 39.

COST PER TON MILE, 39.

COST PER WAGON MILE, 39.

COSTS,
Factory, 215.
Prime, 215.
Running, 39.
Total, 216.

CUMULATIVE ERROR, 55.

CUMULATIVE FREQUENCY, 97, 155.

CURVES,
In Diagrams, 143, 153, 183, 184.
Lorenz, 233.
Normal Frequency, 149.
Of Error, Normal, 149.
Of Normal Tendency, 155.
Smoothing of in Diagrams, 153.

CYCLES OF TRADE, 28, 76.

D

DATA,
Arraying, 96.
Commercial, capable of being accurately compiled, 22
Difficulty of compiling, 22.
Median not suitable for use in, 101.
Tabulation of, 113, 121, 128, 129. 131, 132, 134, 135, 137.
Use of Mode in, 109.
Use of Moving Average in, 19, 77, 90
Comparability of Essential, 11, 28.
Completeness of necessary, 11, 16, 28.
Correctness of important, 11.
Divisions of Statistical, 42.
Homogeneity of, 12.
Information to be obtained from, 39, 301, 311.
Methods of Collection of, 11, 22.
Methods of dealing with, 31.
Methods of presenting, 48, 112, 140.
Presentation of may lead to fallacious conclusions, 11
Published, 26.
Quantities of, give more accurate results, 9, 19. 38, 69, 75
Representative, 19.
Stability of, 19.
Suitable for Tabulation, 121.
Uniformity in collection essential, 11, 22.

DECILES, 99.

DECIMALS,
Advantages of, 67.
Care in use of, when approximating, 67.
In Ratios. care in use of, 67.
Meaning of Correctness to second place of, 49.

DEFINITION OF UNIT,
Care necessary in, 11.
Variation in reading of, 12.

DEFINITIONS,
Aggregate, 70
Arithmetic Average, 70.
Array, 96.
Correlation, 288.
Dispersion, 220.
Error, 53.
Factory Cost, 215
Median, 96.
Mode, 101.
Prime Cost, 215.
Quartile, 98.
Running Cost, 39.
Total Cost, 216.
Weighted Average, 84.

DEGREES OF ACCURACY, How shown, 48.

DELIVERY CHARGES, Records of, 41.

DEMAND AND SUPPLY, Effect of Law of, 40.

DEPARTMENTAL RECORDS, 210.

DESCRIPTIVE AVERAGE, 92.

DEVIATION,
Average, 223.
Quartile, 231.
Standard, 227.

DIAGRAMS,
Areas shown in, 174.
Arrangement of, 182.
Circular, 174.
Curves in, 143, 153, 183, 184.
Examples of, 142, 144, 147, 151, 156, 158, 160, 163, 165, 169, 173, 175, 176, 178, 179, 188, *facing* 196, 198, 202, 206, 209, 234, *facing* 288, 298.
Figures in, 184.
Horizontal Scale of, 184.
Linear Magnitudes most suitable for, 172, 182.
Logarithmic Scale in, 167, 169, 184.
Natural Scale in, 165, 166.
Percentages in, 167, 183.
Rules for Constructing, 182.
Skewness in, 150.
Time element in, 159, 183.
Title of, 184.
Types of—159.
Bar, 162.
Bar showing Components, 171.
Block, 161.
Circular, 176.
Graph, 162.
Vertical Scale of, 161, 182, 184.
Volumes for comparison in, 174.
Zero Lines in, 170, 182, 183.

DISADVANTAGES OF,
Arithmetic Average, 72.
Graphic Methods, 181.
Index Numbers, 242.
Median, 100.
Mode, 107.

DISCRETE SERIES OF DATA, 43.

DISPERSION, 220.
Absolute, 223.
Co-efficient of, 223.
Definition of, 220.
Lorenz Curve, 233.
Measuring, 221.
Modulus, 231.
Moments of, 223.
Quartile Measure of, 231.
Range, 221.
Relative, 223.

DISTRICTS,
Why all must be included in a general inquiry, 21, 305, 307, 311, 327.

E

ECONOMIST, THE
Index Number of, 268.
Sensitive Price Index, 270
Tables from, 124, 125.

EFFECT OF CHANGING THE BASE OF COMPARISON, 13.

ELECTRICAL INDUSTRY,
Costs in, 39.

EMPIRE STATISTICAL BUREAU,
Formation of, 26.
Need for Studying the Returns of, 27.
Objects of, 26

ERROR, 53.
Classes of—
Biassed, 55
Unbiassed, 54.
Definition of, 53.
Effect of on Ratios, 66.
Estimating the, 56.
Examples of, 54.
In an Estimated Average, 59.
Measurement of,
Absolutely, 53.
Relatively, 54.
Normal Curve of, 149.
Probable 294.

ESTATE DUTY,
How Budget, Estimate of, is made. 5.

ESTIMATES,
Accuracy of, 46.
Accuracy of, essential to success in business, 8.
Effect of Natural Factors on, 4.
Examples of in Business, 7, 40.
Reasons why results differ from, 5.
Schedule D Tax, how compiled, 5.
Standard of Accuracy in, 47.
Statistics, the Science of, 6.
Uses of, 3.

ESTIMATING THE QUALITY OF MARKETS, 306.

EXCHANGES,
Fluctuations in Foreign, 41, 200, 203.

EXPENSES,
In Relation to Gross Profit, 133, 203.

EXPORTS,
Increase in the Value of, does not necessarily indicate an Increase in Trade, 17.
Price Indices, 271.
Volume and Value of, 274.

F

FACTORS,
Diagram to show three different, 177.
Measuring the Correlation of, 288.
Natural, Effect of, on Estimates, 4, 242

FACTORY,
Cost, 133, 215.
Output, 18, 32.
Records, 41, 213.

FACTS,
What constitutes Statistical, 43.

FALLACIES IN COMPARISONS, 14, 28, 60.

FALLACIES IN THE USE OF AVERAGES, 35.

FALLACIES IN THE USE OF PERCENTAGES, 60.

FEDERATION OF BRITISH INDUSTRIES, STATISTICAL DEPARTMENT, 28.

FIGURES IN DIAGRAMS, 184.

FINANCIAL RECORDS, 211.

FISHER'S "IDEAL FORMULA," 257.

FLUCTUATIONS,
Abnormal, How to eliminate for the purposes of comparison, 18, 36, 69, 75.
Long Term, 9, 19, 38, 69, 75, 292.
Short Term, 18, 38, 80, 294.

FLUCTUATIONS IN PRICE, 17, 29, 40, 75, 199.

FLUCTUATIONS IN THE RATE OF INTEREST,
Importance to business men of, 41.

FOREIGN EXCHANGES,
Fluctuations in, 41, 200, 203.

FREQUENCY,
Cumulative, 97, 155.
Normal Curve of, 149.
Polygon, 146, 148, 154.
Tables, 96, 150.

FUNCTIONS AND OBJECTS OF BUSINESS RESEARCH, 319.

FUTURES,
Operators in, and their methods, 4.

G

GALTON GRAPH, 298.

GANG HOURS, 32.

GEOMETRIC AVERAGE, 94, 246.
Compared with Arithmetic Mean, 250

GRAPHIC METHOD, THE, 139.
Advantages and Disadvantages of, 181.
Application of, 145.
Applied to Commerce, 185.
Areas not suitable for comparison in, 174.
Linear Magnitudes most suitable for use in, 172, 182.
Percentages in, 167, 183.
Volumes not suitable for comparison in, 177.

GRAPHS, 162.
Galton, 298.
Logarithmic Scale in, 167, 169, 184.
Natural Scale in, 165-166.

GROSS PROFIT,
In relation to Expenses, 133, 203.

GROUPS OF STATISTICAL DATA, 42.

H

HISTOGRAM, 143.

HISTORIGRAM, 161.

HOMOGENEITY,
Importance of, 12.
Necessity for, 13.

HORIZONTAL SCALES IN DIAGRAMS, 184.

I

IMPORTS,
How Estimate of yield of Taxation on is made, 5.
Increase in value of no indication of an increase in Trade, 17.
Price Indices, 271.
Volume and Value, 274.

INCOME TAX,
Method of estimating for Budget purposes, 5.

INDEX NUMBERS, 241.
Average to be used in, 244
Chain Base, 252.
Common uses of, 265.
Correlation of Factors by use of, 289.
Limitations of, 242.
Nature and Principles of, 241.
Relative, the, 243.
Types of—
Board of Trade, 265.
Clothing, Wholesale Stocks and Sales of, 286.
Cost of Living, 275.
Economist, 268.
Economist's Sensitive Price Index, 270.
Import and Export Prices, 271.
Industrial Production, 280.
Interim, 278.

INDEX NUMBERS—(*continued*)
Types of—
London and Cambridge Economic Service, 280.
Retail Prices, Interim of, 277.
Retail Sales, 285.
Sales and Stocks, 285.
Sales of Independent Retailers, 285.
Statist's, 267.
Wages, Rates of, 282.
Wage-rate, Prof Bowley's, 283.
Wholesale Textile Houses, Trade of, 286.
Uses of for purposes of Comparison, 241.

INDUSTRIAL PRODUCTION,
Index of, 280.
Interim Index of, 278.

INERTIA, LAW OF, in large numbers, 20, 38.

INFORMATION,
Obtainable from the Census, 301.
Types of obtainable from Statistical Data, 27, 39, 301, 311.

INSURANCE COMPANIES,
And their use of Statistics, 3, 7, 22

INTEREST,
Fluctuations in the Rate of, 41

INTERPOLATION, 98.

INTERVAL, CLASS, 97.

L

LABOUR SUPPLY,
Information obtainable from Census Returns, 207.

LAG,
How recognised in Diagrams, 190, 291.

LAW OF INERTIA IN LARGE NUMBERS, 20, 38.

LAW OF STATISTICAL REGULARITY IN LARGE NUMBERS, 20

LIMITATIONS OF PUBLISHED DATA, 26.

LINEAR MAGNITUDES IN THE GRAPHIC METHOD, 172, 182.

LIVING, INDEX NUMBER OF THE COST OF, 275.

LOGARITHMIC SCALES,
Use of in Graphs, 167, 169, 184.

LOGARITHMS,
Use of, 67, 168.

LONDON AND CAMBRIDGE ECONOMIC SERVICE,
Index Number of, 280.

LONG TERM OBSERVATIONS,
The importance of, 9, 19, 38, 69, 75, 292.

LORENZ CURVE, 233.

LOWER QUARTILE, 99.

M

MAGNITUDE OF THE MEDIAN,
How to obtain by the use of the Ogive, 157.
Method of calculation, 98.

MAGNITUDES,
Linear only suitable for use in diagrams, 172, 182.

MAN-HOUR,
Advantages of using as a basis of comparison, 32
Example of calculating the Output per, 34.
Output per, 32.

MARKETS,
Estimating the Quality of, 306.
Quantitative factor, 304, 305.

MEAN. (See AVERAGE.)

MEASURES OF DISPERSION, 221.

MEASURES OF ERROR, 53.

MEASURES OF SKEWNESS, 236.

MECHANICAL AIDS TO STATISTICAL WORK, 331

MEDIAN, 96.
Advantages of, 99.
Definition of, 96.
Disadvantages of, 100.
Effect of Skewness on, 152.
Location of, in the Curve of Normal Error, 149.
Not suitable for use in Commercial Work, 101.
Rule for Calculating the Magnitude of, 98.
Rule for obtaining the Magnitude of, from an Ogive, 157.
Uses of, 100.

METHOD, GRAPHIC, 140.
Applied to Commerce, 185.

METHODS OF APPROXIMATION, 48.

METHODS OF COMPARISON, 219.

METHODS OF COMPILING DATA, 11, 22.

METHODS OF PRESENTING DATA, 48, 112, 140.

METHODS OF SAMPLING, 21.

METHODS OF TABULATING, 112.

MODAL OUTPUT, 109.

MODAL TIME, 109.

MODE, 101
Advantages of, 107.
Definition of, 101.
Disadvantages of, 107.
Method of Location of, 102, 159.
Use of, in Commerce, 109.

MODULUS, 231.

MOMENTS OF DISPERSION, 223.

MONEY,
Effect of change of Purchasing Power of, 17, 29, 40, 75.

MONTHLY SALES,
Method of Presenting, 187.

MOVING AVERAGE, 19, 77.
How compiled, 77.
Utility of, 90, 190.

MOVING AVERAGE PRINCIPLE,
Advantages of, 19.
Applied to Commerce, 90.

MOVING WEIGHTED AVERAGE, Uses of, 90.

N

NATURAL FACTORS,
Their Effect on Estimates, 4.

NATURAL SCALE IN GRAPHS, 165, 166.

NET PROFIT AND EXPENSES, 204.

NET PROFIT AND TURNOVER,
Method of Presenting, 208.

NORM. (See MODE.)

NORMAL CURVE OF ERROR, 149.

NORMAL FREQUENCY CURVE, 149.

NUMBERS,
Index, 241.
Law of Inertia in Large, 20, 38.

O

OBJECTS OF BUSINESS RESEARCH, 319.

OBJECTS OF EMPIRE STATISTICAL BUREAU, 26.

OBJECTS OF STATISTICS, 6.

OBSERVATIONS OVER LONG PERIODS,
Necessary to ensure Correctness of Estimates, 9, 19, 38, 69, 75.

OGIVE, 157.
Diagram showing, 158

ONCOST, 216.

OPERATORS IN FUTURES,
Methods of Working, 4.

OUTPUT,
Effect of Bank Holidays on, 18, 32.
Effect of Methods of Production on, 18
Factory, 18, 31, 213.
Modal, 109
Profit Earning and, 31, 115.
Standard of, 18, 32, 41.

OUTPUT PER MAN-HOUR, 32.
Example of Calculation of, 34.

P

PASSENGER MILE,
As a Unit for comparison, 39.

PEARSON, KARL,
Co-efficient of Correlation, 292.

PERCENTAGES,
Effect of Approximation on, 52.
Fallacies in the Use of, 60.
Method of checking calculations of, 130.
Use of,
In Diagrams, 167, 183.
In Railway Returns, 122.
In Tabular Work, 128, 133.

PERCENTILES, 99.

PERIOD OVER WHICH AN AVERAGE SHOULD BE COMPUTED, 76

PICTOGRAM, 172.

PICTORIAL REPRESENTATION,
Advantages of, 140.

POLYGON, Frequency, 146, 148, 154.

POPULATION,
Density of and Market Possibilities, 304.

POWERS SORTING AND TABULATING MACHINES, 332.

PRESENTATION OF DATA,
Importance of Proper, 114.
Methods of, 48, 113, 140.

PRICE,
Fluctuations in, 17, 29, 40, 75, 199.

PRIME COST,
Definition of, 215.

PROBABILITIES,
Business an Application of the Science of, 8.
Theory of, 19.

PROBABILITY OF ERROR,
In co-efficient of correlation, 294.

PRODUCING BUSINESSES,
Factors affecting, 18.
Short Term Observations in, 18.

PRODUCTION,
Organisation of, 322.

PROFITS,
Output of a Factory, no criterion of, 31, 115
What are, 11.
Why returns of for comparison must be compiled on the same basis, 14, 23.

PROGRESSIVE AVERAGE, 81.
Example of, in Commerce, 90.

PUBLISHED STATISTICS,
Limitations of, 26.
Uses of 4, 26.

PURCHASES AND SALES,
Method of Presenting, 196.

PURCHASING POWER OF MONEY,
Effect of Change in, 17, 29, 40, 75.

Q

QUANTITIES,
Comparison of, with values, 194
Necessity for Records of, 39.

QUANTITY, INDEX NUMBERS OF, 260.

QUARTILE MEASURE AND CO-EFFICIENT OF DISPERSION, 231.

QUARTILES, 98.

QUESTIONNAIRE, 23.
Care in the drafting of, 23.
Example of, 25.
Rules for drafting, 23.

QUOTAS,
Examples of use in connection with Selling, 131

R

RAILWAYS,
Example of, Returns of, 119, 122, 127.
Use of Percentages in Returns of, 122.
Use of Statistics by, 4, 7.

RANGE OF DISPERSION, 221.

RATING OF PROPERTY,
Incompleteness of Data relating thereto leads to fallacious conclusions, 18.
Ratio of variation, 295

RATIO OF VARIATION, 295.

RATIOS,
Care of Use of Decimal Point in, 64.
Effect of Approximation on, 66.
Use of, 64.

RECORDS,
Advertising, 41.
Branch, 130.
Control, 211.
Cost, 133, 215.
Departmental, 210
Factory, 41, 213.
Financial, 211.
Necessity for, in Business, 3
Sales, 187.
Transport, 39, 136, 217.

REGRESSION,
Line of, 297.

REGULARITY, Law of Statistical, 20

RELATIVE DISPERSION, 223.

RELATIVE ERROR,
Definition of, 54.
Estimate of the Biassed, 59.
. the Unbiassed, 57.
In an Estimated Average, 60.
Relative Measure of Skewness, 236.

RESEARCH, BUSINESS, 315.
Objects of, 319

RESERVES,
Part of Capital for purposes of Comparison, 16

RETAIL PRICES,
Interim Index of, 277,

RETAIL SALES,
Index of, 285.

RETURNS,
Board of Trade, 17, 26.
Empire Statistical Bureau, 26.
Example of Tabulated, 116, 117, 119, 122, 124, 125, 127, 129, 130, 131, 132, 134, 135, 137.

RULES,
For Calculating the Magnitude of the Median, 98.
For Calculating the Magnitude of the Quartiles, 99.
For Constructing Diagrams, 182.
For Tabulation, 118

RUNNING COST, 39.

S

SALES,
Accumulated, Method of Presenting, 187.
Monthly, Method of Presenting, 188.
Problems in Connection with, 324.
Records of, 187.
Returns, How shown in Diagrams, 192
Tabulated, 129

SALES AND PURCHASES,
Method of Presenting, 196.

SALES AND STOCKS,
Various Indices of, 285

SAMPLING, 20
Methods of, 21.
Uniformity in, 21.
What is meant by, 20.

SCALES IN DIAGRAMS,
Horizontal, 184.
Logarithmic, 167, 169, 184
Natural, 166.
Vertical, 161, 181, 182.

SCHEDULE D ASSESSMENT OF INCOME TAX
How Budget Estimate is made, 5.

SCHEDULE OF QUESTIONS. (See QUESTIONNAIRE)

SELFRIDGE, Mr. H. G.,
On effects of Changes in the Purchasing Power of Money, 31

SERIES OF STATISTICAL DATA, 42.

SHORT TERM FLUCTUATIONS, 18, 38, 80, 294.

SHORT TERM OBSERVATIONS IN A PRODUCING BUSINESS, 38.

SIMPLE ARITHMETIC AVERAGE, 70.

SIMPLE TABULATION, 116.

SKEWNESS, 150.
Example of in Diagram, 151.
How Measured, 236.

SMOOTHING LINES IN DIAGRAMS, 153.

STABILITY OF DATA,
Importance of, 19.
Impossibility of Absolute, 19

STANDARD DEVIATION, 227.
Short Cut Method of Calculating, 228.

STANDARD OF ACCURACY IN ESTIMATES, 47.

STANDARD OF OUTPUT, 18, 32, 41.

STANDARDS FOR COMPARISON,
Moving Average, 9, 19, 75.

STATIST, THE,
Index Number of, 267.

STATISTICAL BUREAU,
Advantages of, 27
Formation of Empire, 26.
Objects of Empire, 26.
Returns of Empire, 26.

STATISTICAL DATA,
Classes, 42.
Information to be obtained from, 39.

STATISTICAL DEPARTMENTS IN BUSINESS, 9, 28, 328.

STATISTICAL REGULARITY, Law of, 20.

STATISTICS,
An Application of the Science of Estimates, 6.
Commerce and, 7.
Commercial Uses of, 40.
Compilation of, 11.
Cost Accountancy a branch of, 7.
Difference between Accountancy and, 7.
Information to be obtained from, 27, 39, 301, 311.
Object of, 6.
Published, 4, 26.
Records of Past Experience, 6.
Use of by Banks, 3, 7.
Insurance Companies, 3, 7, 22.
Railways, 4, 7.
Use of, in Commerce, 7, 39, 301.

STOCK EXCHANGE TRANSACTIONS,
Application of the Progressive Average Principle to, 91.

SUPPLY AND DEMAND, Effect of, Law of, 40.

T

TABLES,
Frequency, 96, 150.

TABULATION, 113.
Arrangement of, 123, 125.
Commercial Data, 113, 121, 128, 131, 132, 134, 135, 137.
Complex, 116
Data for, 121.
Graphic Method of Presentment depends on, 114.
Importance of, 114.
Miscellaneous Items in, 120.
Rules for, 118.
Simple, 116.
Title of Table, 123.
Types of, 116.
Typewriters, Limitations in the use of in, 126.
Suitable for, 128.
Uses of Averages in, 128.
Percentages in, 128.

TAXATION,
How Estimates of yield are compiled, 5.

TENDENCY,
Normal Curve of, 155.

THEORY OF PROBABILITIES, 19

TIME,
How shown in Diagrams, 159, 183.
Modal, 109.

TITLES,
Of Diagrams, 184.
Of Tables, 123.

TON-MILES AS A BASIS OF COMPARISON, 39.

TOTAL COST,
Definition of, 216.

TRADE,
Cycles, 28, 76.
Increase of the Value of Exports no indication of an increase in, 17.

TRANSPORT RECORDS, 39, 136, 217.

TREND, 78, 166, 190.

TURNOVER,
Increase in the value of no indication of Progress of a Business, 29.
Method of showing Net Profit and, 208.
Profits correlated with, 115.

TYPES, 69. (See also AVERAGES.)
Deciles, 99.
Median, 96.
Mode, 101.
Percentiles, 99.
Quartiles, 98.

TYPEWRITERS,
Use of in Tabulation, 126.

U

UNBIASSED ERROR,
Definition of, 54.
Estimating the, 56.

UNIT,
Care necessary in the definition of, 11, 22, 31.
Man-Hour, 32.
Necessity for uniformity in the definition of, 12.
Suitable for Comparison, 12, 31, 39.
When necessary to divide, 13.

UPPER QUARTILE, 99.

V

VALUE,
Of Exports, An increase in, no criterion of increasing business, 17
Of Turnover, An increase in, no indication of the progress of a business, 29.

VALUES AND QUANTITIES,
Comparison of, 194.

VARIATION, RATIO OF, 295.

VERTICAL SCALE IN DIAGRAMS, 161, 182, 184.

VOLUMES,
Care in using for the purposes of Comparison, 177.

W

WAGES,
Example of Workers grouped according to varying rates of, 13.
Index Number of, Rates of, 282.
Per Man-Hour, How calculated, 32.
Professor Bowley's Index Number, 283.

WAGON-MILE,
Cost per, 39.

WEIGHTED AVERAGE, 83.
Application of, Progressive Average to, 90.
Commercial Uses of, 90.
Definition of, 84.
Effect of Changing Proportions of Units, 89.
Weights to be used, 85.
When necessary to use it, 87.

WEIGHTING SYSTEM,
As applied to A.M. and G.M., 254.

WORKERS,
Number of affects Output, 18.

Z

ZERO LINES IN DIAGRAMS, 170, 182, 183.

The Stellar Press, Ltd.,
Union Street,
Barnet, Herts,